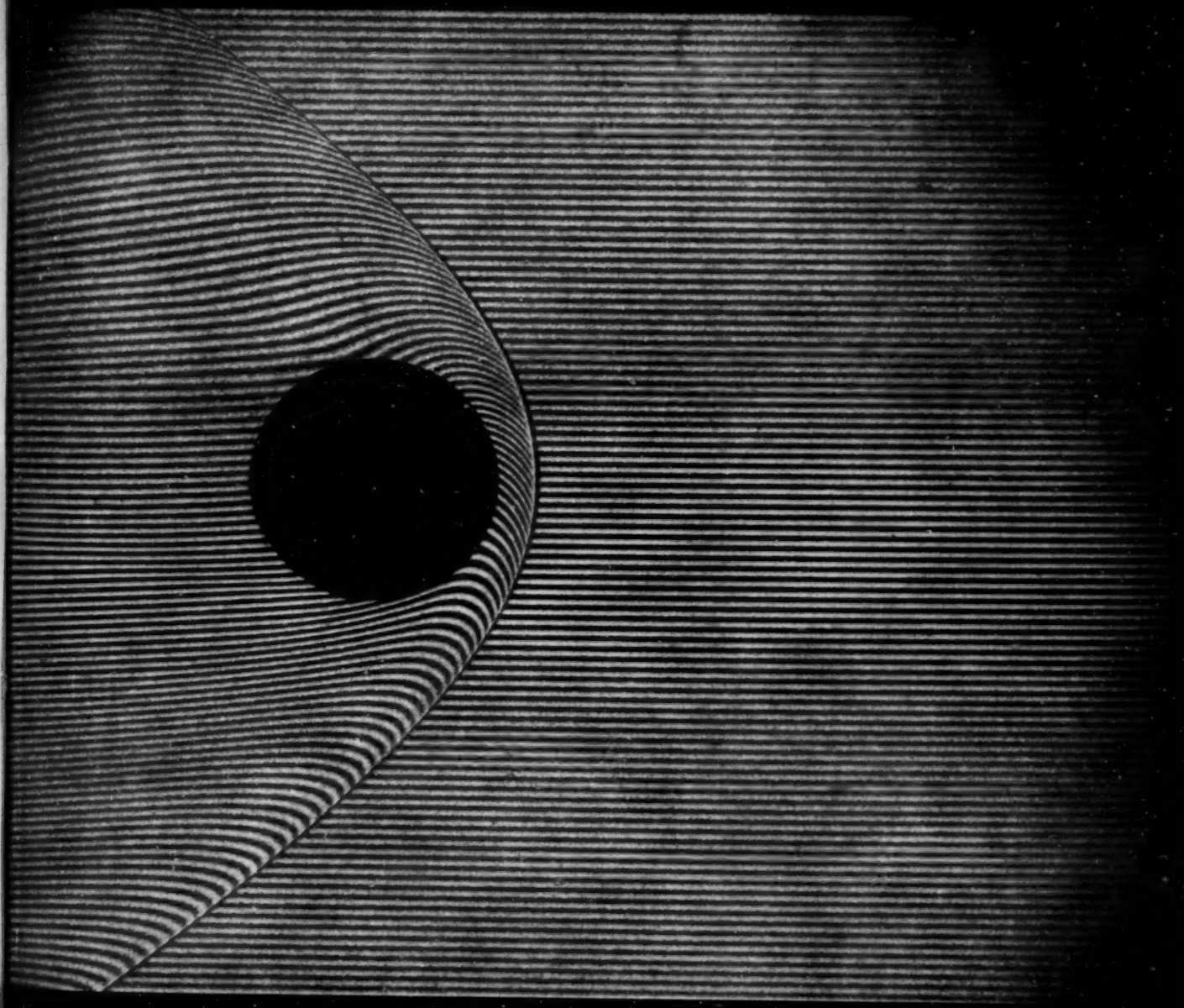
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 23, No. 4 APRIL, 1952

In This Issue

	Page
Statistical Errors in Measurements on Random Time Functions W. B. Davenport, Jr., R. A. Johnson, and D. Middleton	377
Elastic Constants of Nickel J. R. Neighbours, F. W. Bratten, and Charles S. Smith	389
Effect of Thermal-Velocity Spread on the Noise Figure in Traveling-Wave Tubes Philip Parzen	394
Multiple Scattering of Radiation by an Arbitrary Planar Configuration of Parallel Cylinders and by Two Parallel Cylinders Victor Twersky	407
Focusing in a Beta-Ray Spectrometer with Tilted Sources C. M. Fowler and P. Domotor	415
Studies in Newtonian Flow. IV. Viscosity *vs* Molecular Weight in Liquids; Viscosity *vs* Concentration in Polymer Solutions Arthur K. Doolittle	418
Metallurgical Applications of X-Ray Fluorescent Analysis P. K. Koh and Betty Caugherty	427
Deformation and Fracturing of Thick-Walled Steel Cylinders under Explosive Attack John Pearson and John S. Rinehart	434
Influence of Cold-Work on the Hall Effect of a Metal Bernard Wiener and Gerhart Groetzinger	441
Structure of Oxide Replicas for Electron Microscopy Shigeto Yamaguchi	445
Flow over Thick Plate with Circular Hole W. R. Smythe	447
Interferometric Analysis of Airflow about Projectiles in Free Flight F. D. Bennett, W. C. Carter, and V. E. Bergdolt	453
Elastic Constants of Beta-Brass Single Crystals Robert A. Artman and Donald O. Thompson	470
Temperature Dependence of Young's Modulus and Internal Friction of Single Crystals of Beta-Brass Robert A. Artman	475
Radiation Field of a Square, Helical Beam Antenna H. Lottrup Knudsen	483
Steady-State Heat Flow Problem for a Quarter Infinite Solid William Karush	492
Oxygen-Free Single Crystals of Lead Telluride, Selenide, and Sulfide W. D. Lawson	495
Letters to the Editor:	
Separation of Cold-Work Distortion and Particle Size Broadening in X-Ray Patterns B. E. Warren and B. L. Averbach	497
Current-Carrying Wire Loops in a Simple Inhomogeneous Region James R. Wait	497
Study of Oxide-Coated Cathode by X-Ray Diffraction Method Eiso Yamaka	498
Blue Glass Which Fades upon Nuclear Radiation K. O. Otley and W. A. Weyl	499
Effects of Lattice Imperfections on X-Ray Line Broadening M. S. Paterson	499
Transmission of X-Rays at Settings Near the Bragg Angle G. N. Ramachandran	500
Index to Advertisers	xxxvi

BOARD OF EDITORS

Elmer Hutchisson, *Editor*

Earle C. Gregg, *Assistant Editor*

Associate Editors

Howard H. Aiken
Ray E. Bolz
Lyle D. Borst
Robert D. Heidenreich
James Hillier
Leonard B. Loeb
W. James Lyons
Eric Reissner
Roman Smoluchowski

Publication of the American Institute of Physics

OFFICERS

George R. Harrison, *Chairman*
George B. Pegram, *Treasurer*

GOVERNING BOARD

Karl K. Darrow
G. J. Dienes
S. A. Goudsmit
G. P. Harnwell
George R. Harrison
F. V. Hunt
Rudolph Kingslake
Hugh S. Knowles
C. C. Lauritsen
F. W. Loomis
William F. Meggers
George B. Pegram
I. I. Rabi
Duane Roller
Frederick Seitz
Francis G. Slack
John C. Steinberg
Richard M. Sutton
Mark W. Zemansky

Consultant on Printing

Melvin Loos

ADMINISTRATIVE STAFF

Henry A. Barton, *Director*
Wallace Waterfall, *Executive Secretary*
Theodore Vorburger, *Advertising Manager*
Ruth F. Bryans, *Publication Manager*
Betty H. Goodfriend, *Assistant Secretary*
Eileen B. Neuberger, *Circulation Manager*
Kathryn Setze *Assistant Treasurer*

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of **$8** per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $10.00. Single copies: $1.00 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 23, Number 4 April, 1952

Statistical Errors in Measurements on Random Time Functions*

W. B. Davenport, Jr.
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

AND

R. A. Johnson and D. Middleton
Cruft Laboratory, Harvard University, Cambridge, Massachusetts

(Received August 29, 1951)

The statistical errors in time-average measurements of various properties of (stationary) random time functions are analyzed for their dependence on the averaging interval, the method of averaging, and the statistics of the particular time function under observation. The specific averaging processes considered here are as follows:

(1) The continuous time average $\left(\frac{1}{T}\int_0^T Z(t)dt\right)$,

(2) The summation (by discrete sampling techniques) $\left(\frac{1}{N}\sum_{n=1}^{N} Z(nT_0)\right)$, and

(3) The smoothing accomplished by various low-pass filters (for example, a dc meter).

Here $Z(t)$ may represent the output of a linear or nonlinear device. As a particular example, the measurement of the power in a random noise voltage of band width ω_F is analyzed in detail. The expected root-mean-square error approaches asymptotically the value $(K/\omega_F T)^{\frac{1}{2}}$, where K is some numerical constant, depending on the device in question. For (1) and (2) T is the total observation time while for (3) it is proportional to a relevant time constant of the low-pass filter. A short discussion of the relative merits of correlators and spectrum analyzers is included, together with a brief treatment of the distribution function of the error in the measurement.

1. INTRODUCTION

MUCH progress has been made in recent years on the statistical description of signals and random noise in communications systems.[1-8] Theoretical analysis makes it possible to calculate the probability distributions and spectra of random noise at the output of linear filters and nonlinear devices (with resistive loads) provided certain assumptions are made concerning the origin of the noise at the input. Up to now, theory has progressed much more rapidly than experimental technique. This is not surprising, since once certain basic assumptions are made the problems are amenable to mathematical treatment, while the measurement of even the simplest statistic of a random time function requires rather complex equipment and a considerable amount of care and foresight in its application.

The statistical properties of a noise wave, message, or even a periodic wave are contained in an appropriate probability distribution density $W_n(y_1, t_1; y_2, t_2; \cdots y_n, t_n)$, $n \geq 1$, where in general $W_n dy_1 dy_2 \cdots dy_n$ specifies the joint probability that the variable y

* The research reported in this paper was done independently in the Electronics Research Laboratory at Harvard University and the Research Laboratory of Electronics at Massachusetts Institute of Technology. When it was discovered that duplication had occurred, it was decided to publish the results jointly. The work done at Harvard was supported jointly by the ONR, the Signal Corps, and the U. S. Air Force under ONR Contract N5ori-76, T. O. 1. The work done at M.I.T. was supported in part by the Signal Corps, Air Materiel Command, and ONR.

[1] S. O. Rice, Bell System Tech. J. **23**, 282 (1944); **24**, 46 (1945); **27**, 109 (1948).

[2] J. H. Van Vleck and D. Middleton, J. Appl. Phys. **17**, 940 (1946).

[3] D. Middleton, Quart. Appl. Math. **5**, 445 (1948); **7**, 129 (1949); **8**, 59 (1950).

[4] James, Nichols, and Phillips, *Theory of Servomechanisms*, (M.I.T. Radiation Laboratory Series No. 25, 1948) (McGraw-Hill Book Company, Inc., New York, 1947).

[5] N. Wiener, *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (John Wiley and Sons, Inc., New York and Technology Press, New York, 1951).

[6] Lee, Cheatham, and Wiesner, Proc. Inst. Radio Engrs. **38**, 1165 (1950).

[7] D. Middleton, Proc. Inst. Radio Engrs. **36**, 1467 (1948); J. Appl. Phys. **17**, 778 (1946); **20**, 334 (1949).

[8] J. L. Lawson and G. E. Uhlenbeck, *Threshold Signals*, (M.I.T. Radiation Laboratory Series No. 24, 1950) (McGraw-Hill Book Company, Inc., New York, 1950).

(representing the noise, message, etc.) lies in the range (y_1, y_1+dy_1) at time $t=t_1$, in the range (y_2, y_2+dy_2) at $t=t_2\ldots$ and in the range (y_n, y_n+dy_n) at $t=t_n$. Since the underlying mechanism which produces the voltage or current $y(t)$ does not change with time for most problems of interest, the origin of time is of no consequence and the system is said to be stationary. Therefore, we may arbitrarily set any of the times t_1 to t_n (t_1 for convenience) equal to zero and consider only the time differences. For many noise problems the second-order distribution density $W_2(y_1; y_2, t_2)$ gives an adequate statistical description of $y(t)$. More general cases, however, require higher order distributions. Since the totality $(n=1, 2, \cdots\infty)$ of these functions W_n completely describes the statistical behavior of $y(t)$, we should like to be able to measure and calculate them. Difficulties arise because W_n is a function of $2n-1$ parameters (when stationarity is assumed). Practically, this forces us to be content with simpler, less complete statistical descriptions such as the *time* and *ensemble* averages, defined respectively by

$$\langle y_1^{k_1} y_2^{k_2}\cdots y_n^{k_n}\rangle_{Av} = \lim_{T\to\infty}\frac{1}{T}\int_0^T y_1(t)^{k_1} y_2(t+t_2)^{k_2}\cdots y_n(t+t_n)^{k_n}dt$$

and

$$\langle y_1^{k_1} y_2^{k_2}\cdots y_n^{k_n}\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} y_1^{k_1} y_2^{k_2}\cdots y_n^{k_n} W_n(y_1; y_2, t_2; \cdots y_n, t_n)\cdot dy_1 dy_2\cdots dy_n.$$

These moments are equivalent for ergodic processes,† which are the subject of the present investigation. Note that a moment derived from an nth-order distribution is a function of at most $n-1$ parameters. Several of these moments are of special importance: the quantity $\langle y\rangle$, or $\langle y\rangle_{Av}$, represents the mean value or steady component of $y(t)$; $\langle y^2\rangle$, or $\langle y^2\rangle_{Av}$, is the mean square value of y, which is proportional to the power in $y(t)$. Similarly, $\langle y(t_1)y(t_2)\rangle$, or $\langle y(t_1)y(t_2)\rangle_{Av}$, is called the *(auto-)correlation function* of y and is written $R_y(t_2-t_1)$ or $R_y(\tau)$, $(\tau=t_2-t_1)$. The autocorrelation function is particularly important: It contains as special cases $\langle y\rangle$ and $\langle y^2\rangle$, and from it may be determined the power-frequency spectrum, since, by the theorem of Wiener and Khintchine,[9] $R_y(\tau)$ and the spectral density $W(f)$ are each other's (cosine) Fourier transforms. Thus, the measurement or calculation of $R_y(\tau)$ is in principle equivalent to finding the spectrum of $y(t)$.

The theory for calculating the moments at the output of a nonlinear device is in principle complete if the statistics of the input disturbance are known or can be assumed. The need for experimental measurement arises (1) when the statistics of the input are not known, (2) when the mathematical description of the linear or nonlinear device is not adequate, and (3) when, as is frequently the case, the analytic treatment, although possible in principle, is so complex as to be impractical. For instance, the correlation function for voltages representing speech or music cannot be calculated, since explicit statistics (W_2, W_3, etc.) of speech and music are unknown.[10] Similarly, the effect of the actual rounded characteristic of a diode rectifier on the spectral output cannot be determined precisely from the simplified theoretical calculation based on a perfect linear rectifier. The development of experimental techniques capable of measuring moments of time functions accordingly extends the statistical treatment to many problems of interest in communication engineering.

As examples of the experimental approach, several correlators of two general types have been constructed. The first type[11,12] periodically samples the voltage or current wave $y(t)$ and computes the correlation function by averaging a large number of products of pairs of samples which are separated by the time interval τ. The second type[13–15] delays $y(t)$ to obtain $y(t-\tau)$, which is next multiplied by $y(t)$ with the help of some form of analog-multiplier to yield $y(t)y(t-\tau)$. The product is then averaged by an analog-integrator (or a dc meter) to give a measure of the correlation function.

The present paper analyzes the statistical errors introduced in the measurement of time averages when the time available for the measurement is not infinite. Practically, it is desirable to complete the measurement in as short a time as possible, consistent with desired accuracy, in order to minimize the effects of drift in the measuring apparatus and of gradual changes in the characteristics of the process under investigation; in other words, to insure stationarity during the measurement. In general, it will be shown that the mean square error in the measurement of the mean of $Z(t)$, where $Z(t)$ is some linear or nonlinear function of the original variable $y(t)$, depends only on the correlation function of $Z(t)$. Formulas are given relating the mean square error to the number of samples (for the discrete sampling case) and to the averaging time (for the continuous analog device) when the statistics of $Z(t)$ are known.

† In more mathematical language, if an ensemble is ergodic—i.e., stationary, and containing no stationary subsets of probability different from 1 and 0—time and ensemble averages are equal, with probability unity. Then $\langle G(y_1\cdots y_n)\rangle = \langle G[y_1(t+t_1)\cdots y_n(t+t_n)]\rangle_{Av}$, with probability 1. It should be emphasized that the random functions that are considered in practice are *assumed* to be ergodic; actual proof of this appears to be impossible.

[9] N. Wiener, Acta Math. **55**, 117 (1930); A. Khintchine, Math. Ann. **109**, 604 (1934).

[10] W. B. Davenport, "A study of speech probability distributions," M.I.T. Research Laboratory of Electronics, Technical Report No. 148, August, 1950.

[11] H. E. Singleton, Proc. Inst. Radio Engrs. **38**, 1422 (1950).

[12] J. F. Reintjes, "Single-channel correlator," M.I.T. Research Laboratory of Electronics, Quarterly Progress Report, April, 1951.

[13] H. R. Seiwell, Rev. Sci. Instr. **21**, 481 (1950).

[14] A. E. Hastings and J. E. Meade, Naval Research Laboratory Report 3679, June, 1950.

[15] K. N. Stevens, J. Acoust. Soc. Am. **22**, 769 (1950).

When the statistics of $Z(t)$ are not known, the expected mean square error cannot be rigorously obtained, since it depends explicitly on the statistics of Z. However, by evaluating the dependence of the error on system parameters for cases in which $Z(t)$ possesses known statistics, some useful knowledge of the behavior can be gained when only partial information (such as band width) is available concerning $Z(t)$.

Part 2 gives an outline of the theory. In Part 3 the expected mean square error in measuring power is discussed as a function of the observation time for several types of averaging devices, and Part 4 compares correlators and spectrum-analyzers in the light of the preceding results. The paper concludes with two appendices. The first gives a short discussion of the problem of determining the first-order distribution of the error. The second treats the optimum filter response for minimization of the rms error in finite time.

2. STATISTICAL ERRORS

We consider first the statistical errors to be expected when a *finite* time average is used in the measurement of various statistical properties of a random time function $y(t)$, properties which are defined mathematically by infinite time and ensemble averages. Here we examine the time average only, since practical measuring devices work on a single member of the ensemble for a long (but not infinite) time rather than instantaneously on a large number of members of the ensemble. The time average $\langle Z(t)\rangle_{Av}$ of any function of $y(t)$ is defined by

$$\langle Z(t)\rangle_{Av}=\lim_{T\to\infty}\frac{1}{T}\int_0^T Z(t)dt, \tag{2.1}$$

where $Z(t)$ has been derived from the original voltage or current wave $y(t)$ by some previous operation. For example, to measure the mean value $\langle y\rangle_{Av}$, we set $Z(t)\equiv y(t)$. If we wish to determine the mean square value $\langle y^2\rangle_{Av}$, we must first put $y(t)$ through a squaring device and then into the integrator (i.e., we set $Z(t)\equiv y^2(t)$). To find the correlation function for a delay of τ-seconds we set $Z(t)\equiv y(t)y(t-\tau)$. In a similar way, to measure $\int_{y_0}^{\infty} W_1(y)dy$ we could put $y(t)$ through a clipping device which has an output $Z(t)\equiv 1$ for $y(t)>y_0$, and $Z(t)\equiv 0$ for $y(t)<y_0$. From the series of measurements for various values of y_0 we can then obtain the first-order distribution density $W_1(y)$ by numerical differentiation.

To compute the mean square error in a sample mean based on a time average over the finite observation interval, T, we define the sample mean as a random variable according to

$$M(T)=\frac{1}{T}\int_0^T Z(t)dt. \tag{2.2}$$

Following the methods used by Rice and others[1,16,17] in a similar problem we can derive many of the statistical properties of $M(T)$ in terms of the properties of $Z(t)$. Assuming that the operation indicated in Eq. (2.2) is repeated a large number of times, generating an ensemble of values of $M(T)$, we can calculate the (statistical) moments associated with this ensemble as follows:

$$\langle M(T)\rangle=M=\frac{1}{T}\int_0^T\langle Z(t)\rangle dt$$

$$=\langle Z\rangle\frac{1}{T}\int_0^T dt=\langle Z\rangle. \tag{2.3}$$

We have removed $\langle Z(t)\rangle$ from under the integral sign, since $Z(t)$ belongs to a stationary random process, and its mean value is therefore independent of the time of observation. Correspondingly, for the second moment we find

$$\langle M(T)^2\rangle=\frac{1}{T^2}\int_0^T\int_0^T\langle Z(t_1)Z(t_2)\rangle dt_1dt_2,$$

$$=\frac{1}{T^2}\int_0^T\int_0^T R_Z(t_2-t_1)dt_1dt_2, \tag{2.4}$$

$$=\frac{2}{T^2}\int_0^T(T-x)R_Z(x)dx, \tag{2.5}$$

where $R_Z(x)$ is the autocorrelation function of Z.‡ The expected mean square fluctuation in the measurements of $M(T)$ is therefore

$$\sigma_M^2(T)\equiv\langle M^2(T)\rangle-\langle M\rangle^2$$

$$=\frac{2}{T}\int_0^T(1-x/T)\{R_Z(x)-\langle Z\rangle^2\}dx. \tag{2.6}$$

To calculate the expected error in the mean value of $M(T)$, we must know $R_Z(x)$ or calculate it from the statistics of the original wave $y(t)$. Usually, such a calculation requires greater knowledge of the statistics of $y(t)$ than is to be found from the measurement of $\langle Z\rangle$. In other words, to determine *a priori* the expected error in the measurement of a statistic, $\langle Z\rangle$, we must know the more complicated statistic, $R_Z(x)-\langle Z\rangle^2$. The situation is similar to that encountered in the measurement of most physical quantities; we cannot specify the expected error by a theoretical calculation without first knowing the distribution of the quantity under observation. However, certain simple limiting relations may be obtained: For very short observation intervals (i.e., for $T\to 0$) the value of $R_Z(x)$ in (2.6) is very nearly equal to $R_Z(0)$. In this case (2.6) accordingly becomes

$$\sigma_M^2(T)\cong\{R_Z(0)-\langle Z\rangle^2\}\frac{2}{T}\int_0^T\left(1-\frac{x}{T}\right)dx$$

$$=R_Z(0)-\langle Z\rangle^2=\sigma_Z^2. \tag{2.7}$$

[1,16] S. O. Rice, J. Acoust. Soc. Am. **14**, 216 (1943).

[17] See also M. Kac and A. J. F. Siegert, J. Appl. Phys. **18**, 383 (1947).

‡ The step from (2.4) to (2.5) is accomplished by setting $x=t_2-t_1$ and noting that $R_Z(x)$ is always an even function of x.

On the other hand, for very long observation intervals (i.e., for $T\to\infty$) we have from (2.6)

$$\sigma_M^2(T)\cong\frac{2}{T}\int_0^\infty\{R_Z(x)-\langle Z\rangle^2\}dx. \tag{2.8}$$

By the theorem of Wiener and Khintchine, the spectral density $W_{Z-\langle Z\rangle}(f)$ is given by

$$\begin{aligned}W_{Z-\langle Z\rangle}(f)&=4\int_0^\infty R_{Z-\langle Z\rangle}(t)\cos\omega t dt,\quad \omega=2\pi f,\\&=4\int_0^\infty\{R_Z(t)-\langle Z\rangle^2\}\cos\omega t dt.\end{aligned} \tag{2.9}$$

Thus for very long observation intervals we have

$$\sigma_M^2(T)\cong\frac{W_{Z-\langle Z\rangle}(0+)}{2T}, \tag{2.10}$$

and the expected value of the mean square error thus varies inversely with the duration of the observation interval as $T\to\infty$.

Returning to Eq. (2.6), we observe that for a given $R_Z(x)-\langle Z\rangle^2$, $\sigma_M^2(T)$ can be made arbitrarily small by choosing T sufficiently large. Unfortunately, circuits capable of performing the operation, $\int_0^T Z(t)dt$, where Z is a voltage or current, can be built for only limited time intervals. Therefore, we must consider other measures of the mean value.

Costas[18] has derived the expected mean square error in a mean value computed by sampling $Z(t)$ periodically (with period T_0) and then finding

$$M(N,T_0)=\frac{1}{N}\sum_{n=1}^{N}Z_n(nT_0). \tag{2.11}$$

The moments of the distribution of $M(N,T_0)$ can be calculated in the manner outlined in (2.2) to (2.6): namely,

$$\begin{aligned}\sigma_M^2(N,T_0)&=\langle M(N,T_0)^2\rangle-\langle M(N,T_0)\rangle^2\\&=\frac{R_Z(0)-\langle Z\rangle^2}{N}+\frac{2}{N^2}\sum_{K=1}^{N-1}(N-K)\\&\quad\times\{R_Z(KT_0)-\langle Z\rangle^2\},\end{aligned} \tag{2.12}$$

which, as shown by Costas, reduces to our Eq. (2.6) in the limit $(N\to\infty, T_0\to dt)$, such that $NT_0=T$ and $KT_0=x$. Here the expected mean square error is always *larger* for a finite number of samples in the sampling time T than for the continuous sampling as outlined in (2.1)–(2.6). This is to be expected, since discrete sampling does not take advantage of all the information concerning $Z(t)$ that is given in the time T. We also note that for purely random inputs, if the sampling period is much larger than the reciprocal of the mean frequency of $Z(t)$, $R_Z(KT_0)-\langle Z\rangle^2$ is essentially zero and the result reduces to the well-known expression for the variance of the sum of N independent samples, *viz.*, $\sigma_M=\sigma_Z/N^{\frac{1}{2}}$. For inputs containing periodic components with periods commensurable with the sampling period T_0, $R_Z(KT_0)-\langle Z\rangle^2$ does not approach zero for large K. Therefore, σ_M will be essentially independent of N and can be of the order of σ_Z.

Another common way to measure mean values in the laboratory is to use meters or low-pass filters of various types.[19] Thiede[20] has calculated the expected mean-square fluctuation in the reading of a hot-wire ammeter. Also, Fano[21] has investigated the smoothing accomplished by a low-pass RC filter in obtaining correlation functions with an analog machine.

In our present case, the operation of averaging is performed by passing the general input, $Z(t)$, through some low-pass filter. The output, $M(T)$, of the averaging filter may be related to its input by means of a convolution integral involving the filter's unit impulse response, or weighting function, $h(t)$. Specifically,

$$M(T)=\int_0^T h(u)Z(T-u)du, \tag{2.13}$$

where T is the duration of the observation interval. Throughout this paper we assume that the averaging filter is either passive, or absolutely stable, thus requiring that the filter impulse response be zero for negative values of its argument.

At this point, it is convenient to define a new filter impulse response, $h(t,T)$, which is equal to the low-pass filter's unit impulse response throughout the observation interval and is zero thereafter. One has then

$$h(t,T)\equiv\begin{cases}h(t), & (0\le t\le T)\\ 0, & (t>T \text{ and } t<0).\end{cases} \tag{2.14}$$

The corresponding system function, $H(\omega,T)$, is given by the Fourier transform of $h(t,T)$:

$$H(\omega,T)\equiv\int_{-\infty}^{+\infty}h(t,T)e^{-i\omega t}dt. \tag{2.15}$$

The new filter impulse response, $h(t,T)$, (and hence also the system function $H(\omega,T)$) thus combine the effects both of the low-pass averaging filter and of the finite observation interval. In terms of this new impulse response, the averaging filter input-output relation may

[18] J. P. Costas, "Periodic sampling of stationary time series," M.I.T. Research Laboratory of Electronics, Technical Report No. 156, May, 1950.

[19] R. R. Bennett and A. S. Fulton, J. Appl. Phys. **22**, 1187 (1951).

[20] H. Thiede, Elc. Nackr. Tek. **13**, 84 (1936).

[21] R. M. Fano, "Signal-to-noise ratio in correlation detectors," M.I.T. Research Laboratory of Electronics, Technical Report No. 186, February, 1951.

be written

$$M(T)=\int_{-\infty}^{+\infty} h(u, T)Z(T-u)du. \qquad (2.16)$$

Following the steps used in obtaining (2.3) and (2.4), we have for the first and second moments of the sample mean:

$$\langle M(T)\rangle=\int_{-\infty}^{+\infty} h(u, T)\langle Z(T-u)\rangle du$$
$$=\langle Z\rangle\int_{-\infty}^{+\infty} h(u, T)du \qquad (2.17)$$

and

$$\langle M^2(T)\rangle=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(u, T)h(v, T)$$
$$\times\langle Z(T-u)Z(T-v)\rangle dudv$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(u, T)h(v, T)R_Z(u-v)dudv. \qquad (2.18)$$

It therefore follows that

$$\sigma_M^2=\langle M^2(T)\rangle-\langle M(T)\rangle^2$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(u, T)h(v, T)$$
$$\times\{R_Z(u-v)-\langle Z\rangle^2\}dudv. \qquad (2.19)$$

Substituting $x=u-v$ gives

$$\sigma_M^2=\int_{-\infty}^{+\infty} R_h(x)\{R_Z(x)-\langle Z\rangle^2\}dx, \qquad (2.20)$$

where

$$R_h(x)\equiv\int_{-\infty}^{+\infty} h(v, T)h(v+x, T)dv. \qquad (2.21)$$

An interpretation of $R_h(x)$ may be made as follows: Direct application of Parseval's theorem to (2.21) results in

$$R_h(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|H(\omega, T)|^2e^{i\omega x}d\omega$$
$$=\int_0^{\infty} 2|H(\omega, T)|^2\cos\omega x df. \qquad (2.22)$$

Then by considering the Wiener-Khintchine theorem, we see that $R_h(x)$ may be thought of as the autocorrelation function of the output of the averaging filter when the filter input is "white" noise with a spectral density, $W(f)$, of magnitude 2.

The expected value of the relative rms error is defined by the ratio of σ_M to $\langle M\rangle$. From Eqs. (2.16) and (2.13) we accordingly obtain

$$\frac{\sigma_M}{\langle M\rangle}=\frac{\left[\int_{-\infty}^{+\infty} R_h(x)\{R_Z(x)-\langle Z\rangle^2\}dx\right]^{\frac{1}{2}}}{\langle Z\rangle\int_{-\infty}^{+\infty} h(x, T)dx}. \qquad (2.23)$$

The expected value of the relative rms error may also be expressed in terms of the system function $H(\omega, T)$ and the spectral density $W_{Z-\langle Z\rangle}(f)$. We have alternatively,

$$\frac{\sigma_M}{\langle M\rangle}=\frac{\left[\int_0^{\infty}|H(2\pi f, T)|^2W_{Z-\langle Z\rangle}(f)df\right]^{\frac{1}{2}}}{\langle Z\rangle H(0, T)}. \qquad (2.24)$$

Equation (2.24) is obtained by applying Parseval's theorem to the numerator of (2.19) and by using (2.15) in the denominator. As in the case of averaging by integration, we must either know $R_Z(x)$ or calculate it from the statistics of the original wave, $y(t)$, in order to calculate the expected value of the relative rms error. However, certain simple limiting expressions follow as in the previous instance.

For observation intervals of very short duration (i.e., for $T\to 0$), the effective filter response $H(\omega, T)$ becomes wide compared to the spectral density $W_{Z-\langle Z\rangle}(f)$. In this case, therefore,

$$|H(2\pi f, T)|^2\cong H^2(0, T) \qquad (2.25)$$

over the range of significantly nonzero values of $W_{Z-\langle Z\rangle}(f)$. Using the result in (2.24) we obtain

$$\frac{\sigma_M}{\langle M\rangle}\cong\frac{\left[\int_0^{\infty} W_{Z-\langle Z\rangle}(f)df\right]^{\frac{1}{2}}}{\langle Z\rangle}$$
$$=\frac{[R_Z(0)-\langle Z\rangle^2]^{\frac{1}{2}}}{\langle Z\rangle}=\frac{\sigma_Z}{\langle Z\rangle}. \qquad (2.26)$$

At the opposite extreme, for very long observation intervals (i.e., $T\to\infty$), the effective band width of the averaging filter is very narrow compared to the effective width of $W_{Z-\langle Z\rangle}(f)$. By our assumption that $T\to\infty$, we have

$$h(t, T)\cong h(t), \qquad (2.27)$$

and consequently,

$$H(\omega, T)\cong H(\omega). \qquad (2.28)$$

By virtue of the assumption concerning the relative band widths, we therefore find that

$$W_{Z-\langle Z\rangle}(f)\cong W_{Z-\langle Z\rangle}(0+) \qquad (2.29)$$

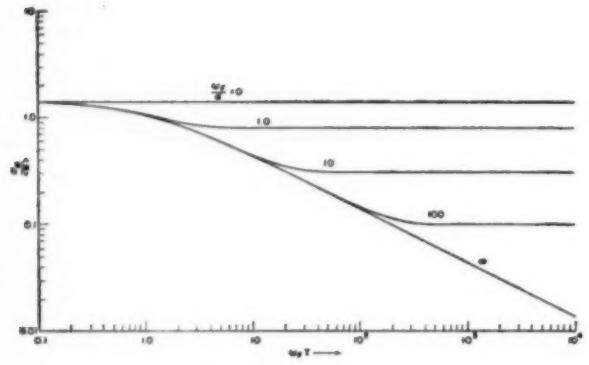


FIG. 1. The relative rms error in measuring the total power in a noise wave of band width, ω_F, (see text) as a function of the averaging time T. The averaging device is a low-pass filter with $RC=1/\alpha$. The corresponding curve for an ideal integrator is identical to that for $\omega_F/\alpha=\infty$.

over the range of significantly nonzero values of $H(\omega, T)$. Using these results in (2.24) we then obtain

$$\frac{\sigma_M}{\langle M\rangle}\cong\frac{\left[W_{Z-\langle Z\rangle}(0+)\int_0^\infty|H(2\pi f)|^2df\right]^{\frac{1}{2}}}{\langle Z\rangle H(0)}. \tag{2.30}$$

If we now define the effective band width, B_h, of the averaging filter as

$$B_h\equiv\frac{\int_0^\infty|H(2\pi f)|^2df}{H^2(0)}, \tag{2.31}$$

it follows from (2.30) that

$$\frac{\sigma_M}{\langle M\rangle}\cong\frac{[W_{Z-\langle Z\rangle}(0+)B_h]^{\frac{1}{2}}}{\langle Z\rangle} \tag{2.32}$$

as the limiting expression for the expected value of the relative rms error in the case of very long observation intervals and very narrow filter band widths. Thus it is seen that in the case of *fixed* filter characteristics, the expected error approaches a constant value as the observation time is increased indefinitely.

However, under certain conditions, the error to be expected may be less than that indicated. For example, in practice a meter can be read more accurately than the mathematics indicates, since an observer will automatically average the reading over a period of time and thus read closer to the true mean than the meter indicates at any particular instant. If a recording meter is used, this "second averaging" can be performed from the recorded $M(T)$.

3. APPLICATION TO PARTICULAR MEASUREMENTS

As a specific application of the expressions developed in Part 2, we now calculate the expected relative rms error in the measurement of the mean power in a normal random noise voltage, $y(t)$, at the output of a low-pass RC filter.[19] To accomplish the measurement we square $y(t)$ with an analog-multiplier and pass the resulting $Z(t)=y^2(t)$ into our averaging device. The relevant correlation functions are

$$\begin{aligned}R_y(\tau)&=R_y(0)e^{-\omega_F|\tau|},\\ R_Z(\tau)&=R_y^2(0)\{1+2e^{-2\omega_F|\tau|}\},\\ (R_Z(\tau)-\langle Z\rangle^2)/\langle Z\rangle^2&=2e^{-2\omega_F|\tau|},\end{aligned} \tag{3.1}$$

where $\omega_F=1/RC$ is the band width (to the half-power point) of the original wave, $y(t)$. If we use for the averaging device a simple RC filter with impulse response (for finite time measurements),

$$h(u, T)=\begin{cases}\alpha e^{-\alpha u} & 0\le u\le T\\ 0 & u<0,\quad u>T,\end{cases} \tag{3.2}$$

the expected value of the relative rms error is then evaluated by substitution in Eq. (2.19):

$$\frac{\sigma_M}{\langle M\rangle}=\frac{\left[2\int_{-\infty}^\infty R_h(x)e^{-2\omega_F|x|}dx\right]^{\frac{1}{2}}}{\int_{-\infty}^\infty h(u, T)du}. \tag{3.3}$$

The denominator is simply

$$\begin{aligned}\int_{-\infty}^\infty h(u, T)du&=\alpha\int_0^T e^{-\alpha u}du\\ &=1-e^{-\alpha T}.\end{aligned} \tag{3.4}$$

The function $R_h(x)$ is determined by substitution in Eq. (2.17). Changing the limits of integration to show explicitly the finite range over which $h(u, T)$ is nonvanishing, we have

$$R_h(x)=\begin{cases}0, & x>T\\ \displaystyle\int_0^{T-x}h(v)h(v+x)dv, & 0\le x\le T\\ \displaystyle\int_{-x}^{T}h(v)h(v+x)dv, & -T\le x\le 0\\ 0, & x<-T.\end{cases} \tag{3.5}$$

Evaluating this expression, with $h(u)=\alpha e^{-\alpha u}$, yields

$$R_h(x)=\begin{cases}\dfrac{\alpha}{2}e^{-\alpha|x|}[1-e^{-2\alpha T}e^{+2\alpha|x|}], & 0\le|x|\le T\\ 0, & |x|>T.\end{cases} \tag{3.6}$$

By substituting the foregoing expressions into Eq. (3.3) and performing the indicated integration over x, we

have for the expected value of the relative rms error

$$\frac{\sigma_M}{\langle M\rangle}=\left[\frac{1}{\frac{1}{2}+(\omega_F/\alpha)}+\frac{e^{-2\alpha T}}{\frac{1}{2}-(\omega_F/\alpha)}-\frac{1}{\frac{1}{2}-(\omega_F^2/\alpha^2)}e^{-(\alpha+2\omega_F)T}\right]^{\frac{1}{2}}\Big/(1-e^{-\alpha T}). \quad (3.7)$$

Obviously, the error depends only on the ratio of the sample length T to some relevant time constants of the input wave and of the averaging filter. Therefore, in Fig. 1 we have plotted $\sigma_M/\langle M\rangle$ as a function of the dimensionless integration "time" $\omega_F T$ with ω_F/α, the ratio of the band width of the input wave to the band width of the averaging filter, as a parameter.

If we use an *ideal* integrator, for which

$$h(u, T)=\begin{cases}1, & 0\leq u\leq T\\ 0, & u<0,\quad u>T,\end{cases} \quad (3.8)$$

as a smoothing device, we obtain (for the same input)

$$R_h(x)=\begin{cases}T-|x|, & 0\leq|x|\leq T\\ 0, & |x|>T,\end{cases} \quad (3.9)$$

$$\int_0^T h(u, T)du=T$$

and

$$\frac{\sigma_M}{\langle M\rangle}=\frac{1}{\omega_F T}[2\omega_F T-1+e^{-2\omega_F T}]^{\frac{1}{2}}. \quad (3.10)$$

This result is also shown in Fig. 1.

In Appendix B the relative rms error for the *optimum* filter in this instance is found to be

$$\frac{\sigma_M}{\langle M\rangle}=\left[\frac{2}{1+\omega_F T}\right]^{\frac{1}{2}}. \quad (3.11)$$

Note that in this example the rms error obtained by the ideal integrator is only slightly greater than that for the optimum filter, particularly in the region of practical interest when $\sigma_M/\langle M\rangle$ is of the order of 10 percent or less (see Fig. 4).

Equation (2.32) is convenient for comparing the expected value of the rms error of several types of low-pass filters in the limit of very long observation intervals and very narrow filter band widths. For the problem discussed above $W_{Z-\langle Z\rangle}(0+)=4\langle Z\rangle^2/\omega_F$, from the transform of $R_{Z-\langle Z\rangle}(\tau)$ (see Eq. (3.1)). Inserting this in (2.32) gives

$$\frac{\sigma_M}{\langle M\rangle}=2\left[\frac{B_h}{\omega_F}\right]^{\frac{1}{2}}. \quad (3.12)$$

The effective band widths (B_h) for several filters are included in Table I. This limiting expression is sufficiently accurate for design use in the practical cases where the relative rms error is desired to be of the order of 10 percent or less.

TABLE I. The effective band width, B_h, of several common types of low-pass filters as defined by Eq. (2.31).

Filter	$h(t)$	B_h
Integrator	$1/T,\ 0<t<T$	$1/2T$
RC filter	$\alpha e^{-\alpha t}$	$\alpha/4$
Critically damped meter	$\gamma^2 t e^{-\gamma t}$	$\gamma/2$
Underdamped meter	$\left(\frac{\beta^2+\omega_0^2}{\omega_0}\right)e^{-\beta t}\sin\omega_0 t$	$\frac{\beta^2+\omega_0^2}{8\omega_0}$
Overdamped meter	$\frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}(e^{-\lambda_1 t}-e^{-\lambda_2 t})$	$\frac{\lambda_1\lambda_2}{4(\lambda_2+\lambda_1)}$

4. CORRELATORS AND SPECTRUM-ANALYZERS

We now turn to a more detailed analysis of the measurement of the autocorrelation function and the power-frequency spectrum. Phillips[22] shows that if the spectrum is to be obtained numerically from recorded data, it is often less work to calculate the correlation function and perform a numerical Fourier transform, than it is to calculate the spectrum directly from the record by a direct Fourier expansion.

Tukey[23] has analyzed the propagation of error in taking transforms of experimentally determined correlation functions and finds Phillips' treatment too simplified. In any case, the problem of whether to measure the correlation function and take its transform, or to make a direct measurement of the spectral density, is not easily resolved. Here, the expected error in a specific measurement as a function of the averaging time is discussed for both methods. We have used the minimum time required for the measurement as a criterion for choosing between the two approaches in order to minimize the inevitable instability of the measuring equipment. Before comparison is made, brief descriptions of the methods used to obtain correlation functions and spectra are included.

Several correlators of two general types have been constructed in recent years. The first type[11,12] samples the voltage or current wave $y(t)$ periodically. From Eq. (2.12) we see that the correlation function measured in this way approaches the mathematically infinite time average as N (for fixed sampling period) is increased without limit, provided there are no components of $y(t)$ which have the same period at which the samples are taken. The second type of measuring device[13–15] first delays $y(t)$ to obtain $y(t-\tau)$, and then multiplies this by

[22] James, Nichols, and Phillips, *Theory of Servomechanisms*, M.I.T. Radiation Laboratory Series No. 25 (McGraw-Hill Book Company, Inc., New York, 1947), pp. 284 *et seq.*

[23] J. W. Tukey, "The sampling theory of power spectrum estimates," "Symposium on applications of autocorrelation analysis to physical problems," June, 1949. (Published by the ONR, May, 1950.)

$y(t)$ in an analog-multiplier to obtain $y(t)y(t-\tau)$. The resulting product is next averaged by an analog-integrator (or a dc meter) to provide a measure of the correlation for a delay time τ. As the integration time (or the meter time constant) is increased without limit, the measurement approaches the mathematically infinite time average. Equation (2.6) or (2.19) gives the expected mean square error as a function of the integration time T [or the meter response $h(t)$]. In either case the error is found to depend on the explicit form of $R_Z(x)-\langle Z\rangle^2=\langle Z(t)Z(t+x)\rangle-\langle Z\rangle^2$, where now $Z(t)=y(t)y(t+\tau)$, since we are measuring a correlation function. Therefore, to evaluate the expected mean-square error we must know (or assume) the form of $\langle y(t)y(t+\tau)y(t+x)y(t+\tau+x)\rangle-\langle y(t)y(t+\tau)\rangle^2$, which is a moment from the fourth-order distribution of the input variable. Since this moment is not known in general, we must resort to evaluations for which the specific form of the fourth moment is given and then use this information to estimate errors when the exact statistics of the input variable are unknown. For the special cases of normal random noise and periodic signals the expected (absolute) error in the measurement of $R(\tau)$ is found to be less than or equal to the expected error in the measurement of $R(0)$. For more general inputs (such as speech or music) this is not known to be true, although it seems reasonable that the expected (absolute) mean square error in the measurement of $R(\tau)$ should not change by orders of magnitude as τ is varied (when the integration time is maintained constant).

To determine experimentally the power-frequency spectrum, $y(t)$ is fed into an appropriate selective circuit. The output is connected to a mean square meter which reads the average power in the narrow band of frequencies about the central frequency of the filter. As the band width is decreased indefinitely and the averaging time (for the meter) is increased, the measurement yields in the limit the mathematically defined quantity $W(f)df$. The selective circuit may be one of the following:

1. A tuned circuit which is moved along the frequency axis by changing the values of its components.
2. A heterodyne system[24] in which the frequencies of the original wave $y(t)$ are mixed with a tunable, strong local oscillator in a frequency-converter, fed into a fixed tuned circuit, amplified, detected, and then averaged.
3. A feedback amplifier in which the feedback loop contains a tuned circuit that is highly degenerative at all but a narrow band of frequencies. Changing the tuning of the feedback loop thus changes the frequencies that are not attenuated in the amplifier.

The first and third methods employ a filter which must be tuned in order to vary the frequency f in $W(f)$. Since the other properties of the filter, such as the midband attenuation and band width, may change with tuning, the measured value of $W(f)$ is subject to errors that are not easily compensated. With the second method, the filter is not changed, since the local oscillator frequency determines what part of the spectrum we are "looking" at.

The mean square value of the fluctuations in the reading of a spectrum-analyzer can be obtained from the results of Part II;§ $Z(t)$ is set equal to the square of the filtered wave $y(t)$. Once again, analytic answers follow only for periodic waves and for those cases in which $y(t)$ has a specified statistical structure.

The following example indicates the main features to be considered when one decides whether to use correlation or spectral analysis for a particular problem: Let us suppose that $y(t)$ is a normal random noise voltage at the output of an $i-f$ stage. We wish to determine experimentally the spectral distribution of $y(t)$ with the help of a correlator or a spectrum-analyzer. The band width of the filter in the analyzer must be the order of one-tenth (or less) the band width of the input $y(t)$ for suitable definition. Assuming here that the filter is a single-tuned circuit and that the input spectrum is flat over the narrow band, we easily find the correlation function of the output of the filter to be $R_y(0)e^{-\omega_A|\tau|}\cos\omega_0\tau$, where $\omega_A/2\pi$ is the band width of the filter and $\omega_0/2\pi$ is its center frequency. The correlation function of the square of the output of the filter is

$$R_Z(\tau)=R_y^2(0)(1+2e^{-2\omega_A|\tau|}\cos^2\omega_0\tau)$$
$$=R_y^2(0)[1+e^{-2\omega_A|\tau|}+e^{-2\omega_A|\tau|}\cos 2\omega_0\tau].$$

The first term is the desired total average power passed by the filter; the second is the low frequency continuum which causes the fluctuation in the reading of the analyzer. The third term represents power at a high frequency, which is easily removed by filtering and may be neglected. We note that the rms error (as a function of the averaging time T_A and the band width ω_A) is given by Fig. 1 (except for a factor $1/\sqrt{2}$), since $R_Z(\tau)-\bar{Z}^2$ in the expression for σ^2 is the same here as that used in the derivation of Table I, except for a factor $\frac{1}{2}$. (The factor $\frac{1}{2}$ arises because in this present instance the low frequency spectrum contains only half of the total fluctuation.) We therefore find that the rms error for each point on the measured spectrum is equal to some numerical constant divided by $(\omega_A T_A)^{\frac{1}{2}}$. If now we wish to increase the definition by using a still narrower filter and taking more points, we must also increase the averaging time, T_A, in order to keep the fluctuations in our measurements at the same relative amplitude. Thus, to double the number of points and double the

[24] J. D. Cobine and J. R. Curry, Rev. Sci. Instr. **17**, 190 (1946).

§ If the frequency-selective element in the analyzer is swept along the frequency axis too rapidly, an additional error, which may be of considerable amplitude, will be introduced. This error can be calculated from the rate of change of $W(f)$, the sweeping speed and the impulse response of the recording device. The above analysis gives only the statistical error in determining $W(f)$ for fixed f. Similarly, there is an additional error in analog-correlators in which $R(\tau)$ is recorded continuously as τ is varied continuously.

resolution, four times the original total time of measurement is needed.

If a correlator is used instead, we find that the rms error in $R_y(0)$ depends on $\omega_F T_C$, where now ω_F is the total band width of the $i-f$ strip and T_C is the effective averaging time of the correlator for a single observation. The correlation function $R_y(\tau)$ has the form $R_y(0)\rho(\tau)\cos\omega_0\tau$, where $\rho(0)$ is unity and ρ is a slowly varying function of τ. If we assume that the error in the measured $\rho(\tau)$ is of the same order of magnitude as the error in $\rho(0)$, the total time necessary for the measurement is equal to the number of points of $\rho(\tau)$ that are required, multiplied by T_C. To double the resolution effectively, twice as many points of $\rho(\tau)$ are needed. These are obtainable (with the same accuracy) in twice the original total time, since the necessary smoothing time per point remains unchanged. Note that the total time required for a measurement of the entire correlation function, $R_y(\tau)$, with a specified accuracy varies as N/ω_F, where N is the number of points and ω_F is the band width of $y(t)$. The total time needed for the measurement of M points on the corresponding spectrum $W(f)$ requires a total time of the order of M^2/ω_F, if we assume the analyzer's filter width ω_A to be comparable to ω_F/M.

The preceding discussion is not intended as a rigorous exposition of the advantages of correlation analysis over spectrum analysis but serves merely to indicate some of the general arguments in favor of the correlator vis-à-vis the spectrum-analyzer. It is not clear, however, how the errors in a measured (or numerically computed) correlation function are propagated when its Fourier transform is taken; nor is it certain that N points describing a correlation function give as much information as the same number of points describing a spectrum. The apparent saving in time, mentioned in the preceding paragraph when a correlator is used, is related to the fact that the correlator operates on *all* the information concerning $y(t)$ that is available during a given observation time, while the spectrum-analyzer filters out only a small portion of the information (that contained in the band width of the analyzer, which is necessarily much smaller than the total band width). Finally, it should be remarked that the criterion by which we distinguish the two methods is not suitable for every problem. Frequently we are concerned less with the amount of time necessary to make the measurement (although this has been found to be an important factor in machines which have been built) than with the feasibility and simplicity of the equipment.

APPENDIX A. FURTHER STATISTICAL PROPERTIES OF THE SAMPLE MEAN

It is of interest to inquire further as to the statistical properties of the sample mean as a function of the observation time T, the statistics of the input $Z(t)$, and the type of measuring device employed. Applying the methods of Part II, we can derive the correlation function of $M(T)$ as follows:

$$\begin{aligned} R_M(\tau) &= \langle M(T)M(T+\tau)\rangle \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(u,T)h(v,T)\langle Z(T-u)Z(T+\tau-v)\rangle du\,dv \\ &= \int_{-\infty}^{\infty} R_h(x)R_Z(\tau+x)dx, \end{aligned} \tag{A.1}$$

in the notation of Part II. Clearly all the higher moments of M can be evaluated in terms of the moments of $Z(t)$. For the particular measurement of noise power described in Part III [$Z(t)=y^2(t)$], the moment $R_Z(\tau+x)$ can be determined from the normal distribution of $y(t)$, yielding

$$\left.\begin{aligned} R_Z(\tau+x) &= \langle y^2\rangle^2\{1+2e^{-2\omega_F|\tau+x|}\} \\ \text{and} \\ R_M(\tau) &= \langle Z\rangle^2\int_{-\infty}^{\infty} R_h(x)\{1+2e^{-2\omega_F|\tau+x|}\}dx \end{aligned}\right\}. \tag{A.2}$$

Since the statistical behavior of a random variable is completely described by the distributions of the variable, a discussion of the distributions of the sample mean is of some interest. Unfortunately, the calculation of these distributions in closed form seems impossible in the general case, as is apparent from the following discussion of the first-order distribution of $M(T)$.

The method of characteristic functions is the straightforward approach to the problem: Let $F_1(\xi)_M$ be the characteristic function of M, which is related to the distribution density $W_1(M)$ of M by

$$W_1(M) = \int_{-\infty}^{\infty} e^{-i\xi M}F_1(\xi)_M\frac{d\xi}{2\pi}$$

$$\begin{aligned} F_1(\xi)_M = \langle e^{i\xi M}\rangle &= \int_{-\infty}^{\infty} e^{i\xi M}W_1(M)dM \\ &= \int_{-\infty}^{\infty}\exp\left[i\xi\int_{-\infty}^{\infty} Z(T-u)h(u,T)du\right]W_1(Z)dZ \end{aligned} \tag{A.3}$$

$$= \sum_{n=0}^{\infty}\frac{(i\xi)^n}{n!}\left\langle\left[\int_{-\infty}^{\infty} Z(T-u)h(u,T)du\right]^n\right\rangle. \tag{A.4}$$

The various moments can be calculated from

$$\langle M^K\rangle = (-i)^K\frac{\partial^K}{\partial\xi^K}F_M(\xi)\bigg|_{\xi=0}, \tag{A.5}$$

which incidentally leads to the formulas for $\langle M\rangle$ and $\langle M^2\rangle$ already derived in Part II. The coefficient of $(i\xi)^n/n!$ in (A.4) is, in fact, the nth moment, $\langle M^n\rangle$, which can be rewritten

$$\begin{aligned} &\left\langle\left[\int_{-\infty}^{\infty} Z(T-u)h(u,T)du\right]^n\right\rangle \\ &\quad= \int_{-\infty}^{\infty} du_1h(u_1,T)\int_{-\infty}^{\infty} du_2h(u_2,T)\cdots \\ &\quad\times\int_{-\infty}^{\infty} du_nh(u_n,T)\langle Z(T-u_1)Z(T-u_2)\cdots Z(T-u_n)\rangle. \end{aligned} \tag{A.6}$$

Therefore, if the n-fold distribution of Z is known or assumed, the indicated multiple average $\langle Z(T-u_1)Z(T-u_2)\cdots Z(T-u_n)\rangle$ is obtained and the n-fold integration performed. Consider first the particular case in which the input to the meter is shot or thermal noise which has previously been passed through a linear filter. Then it can be shown that $Z(t)$ belongs to a stationary random process with normal statistics. For this particular example (Z distributed normally) the present approach is unnecessarily complicated. Since the meter is merely a second linear filter and the input is distributed normally, the output is distributed normally, with $\langle M\rangle$ and $\langle M^2\rangle$ as calculated in Part II. However, in the more interesting cases (such as the measurement of power spectra or correlation functions) the input variable $Z(t)$ is not normally distributed and the explicit evaluation of the distribution, $W_1(M)$, is not possible in general. The approach outlined in (A.3)–(A.6) remains valid, nevertheless, because no specific assumptions as to the statistics of $Z(t)$ have been made. However, the evaluation of these moments and integrals is far more difficult and in general not possible in closed form.

As an example, consider the calculation of $W_1(M)$ by the method outlined in (A.3) to (A.6), when we attempt to measure the mean

power of a normal random noise voltage or current. We then set $Z=y^2$ and the moments

$$\langle Z(T-u_1)Z(T-u_2)\cdots Z(T-u_n)\rangle = \langle y^2(T-u_1)y^2(T-u_2)\cdots y^2(T-u_n)\rangle$$

can be derived from the n-fold multivariate normal distribution of $y(t)$. The integrals indicated in (A.6) are carried through (with difficulty) and the moments evaluated. Since the resulting infinite series for $F_1(\xi)_M$ usually cannot be summed, however, we are unable to find $W_1(M)$ explicitly.

A physical approach to the problem of measuring the mean square of $y(t)$ is perhaps more enlightening. Consider the expression for $M(T)$:

$$M(T)=\alpha\int_0^T y^2(T-u)e^{-\alpha u}du,$$

where we have assumed the smoothing filter to be a low-pass RC network with $RC=1/\alpha$. This can be rewritten as

$$M(T)=\alpha\sum_{n=1}^{n=T/T_0}\int_{(n-1)T_0}^{nT_0} e^{-\alpha u}y^2(T-u)du \qquad (A.7)$$

if the intervals T_0 are long compared to the reciprocal of the band width of the frequencies of $y(t)$. Then if α is small enough so that a large number, N, of the terms of the summation are of the same order of magnitude, $M(T)$ is the sum of N essentially independent samples. Under these conditions the central limit theorem of probability[25] tells us that the distribution of the standardized variable $m=(M-\langle M\rangle)/\sigma_M$ approaches a stationary normal distribution of the form $W_1(m)=[1/(2\pi)^{\frac{1}{2}}]e^{-m^2/2}$ as $N\to\infty$, or, correspondingly $\alpha\to 0$.

We are interested in the distribution when α is very small, but not actually zero. Since the moments can be evaluated in terms of α and the statistics of the input wave, the distribution $W_1(m)$ is next developed in an Edgeworth series[26] of the form,

$$W_1(M)\cong\frac{1}{\sigma_M}\left\{\phi^{(0)}(m)-\frac{1}{3!}\frac{\mu_3}{\mu_2^{\frac{3}{2}}}\phi^{(3)}(m) +\left[\frac{1}{4!}\left(\frac{\mu_4}{\mu_2^2}-3\right)\phi^{(4)}(m)+\frac{10}{6!}\frac{\mu_3^2}{\mu_2^3}\phi^{(6)}(m)\right]+\cdots\right\}. \qquad (A.8)$$

Here μ_K is the Kth central moment $\langle(M-\langle M\rangle)^K\rangle$, and $\phi^{(K)}(x)=(d^K/dx^K)\phi^{(0)}(x)=(-1)^K H_K(x)\phi^{(0)}(x)$ where $H_K(x)$ is the Hermite polynomial of degree K and $\phi^{(0)}(x)$ is the normal distribution, $[1/(2\pi)^{\frac{1}{2}}]e^{-x^2/2}$. The coefficients have been evaluated in the case where we measure the mean power associated with a normal random noise voltage $y(t)$, which has the correlation function $R_y(\tau)=R_y(0)e^{-\omega_F|\tau|}$ corresponding to the spectrum of a thermal noise passed through a low-pass RC network with $RC=1/\omega_F$. The coefficients up to and including those involving the fourth moment of the distribution (for $\omega_F T\gg 1$, i.e., after the transient has decayed) are given by

$$\left.\begin{aligned}\frac{1}{3!}\frac{\mu_3}{\mu_2^{\frac{3}{2}}}&=\frac{\sqrt{2}}{3}\frac{(1+2\omega_F/\alpha)^{\frac{3}{2}}}{1+\omega_F/\alpha}=0(\omega_F/\alpha)^{-\frac{1}{2}},\\ \frac{1}{4!}\left(\frac{\mu_4}{\mu_2^2}-3\right)&=\frac{1}{2}\frac{3+5\omega_F/\alpha}{(1+\omega_F/\alpha)(3+2\omega_F/\alpha)}\\ \frac{10}{6!}\left(\frac{\mu_3^2}{\mu_2^3}\right)&=\frac{1}{9}\frac{1+2\omega_F/\alpha}{(1+\omega_F/\alpha)^2}\end{aligned}\right\}=0(\omega_F/\alpha)^{-1}. \qquad (A.9)$$

Inspection of these terms shows that as the ratio, β, of the band width (ω_F) of the input to the band width (α) of the filter, is made very large, e.g., $\beta\to\infty$, the distribution of the standardized variable $m=(M-\langle M\rangle)/\sigma_M$ approaches the normal distribution $\phi^{(0)}(m)$. For sufficiently large values of $\beta=\omega_F/\alpha$, Eq. (A.8) can be reduced specifically to

$$W_1(M)\cong\frac{1}{\sigma_M}\left\{\phi^{(0)}(m)-\frac{2}{3\beta^{\frac{1}{2}}}\phi^{(3)}(m) +\frac{1}{\beta}\left[\frac{5}{4}\phi^{(4)}(m)+\frac{1}{9}\phi^{(6)}(m)\right]+\cdots\right\}. \qquad (A.10)$$

A typical distribution is shown in Fig. 2.

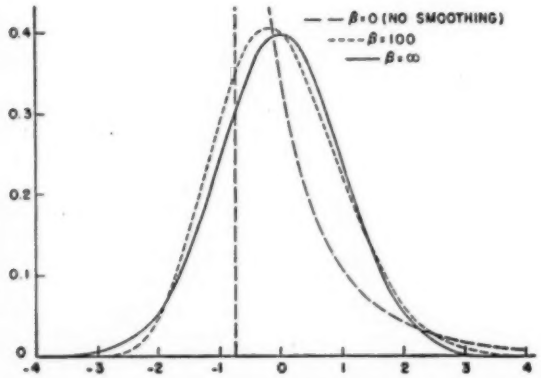


FIG. 2. The distribution of the standardized variable $(M-\langle M\rangle)/\sigma_M$, showing the approach to the normal distribution as $\beta\to\infty$. $M=\alpha\int_0^\infty Z(t-u)e^{-\alpha u}du$, where $Z(t)=y^2(t)$ is the square of the normal random noise passed by a low-pass RC filter with $RC=\omega_F$. $\beta=\omega_F/\alpha$ is the ratio of the band width (ω_F) of the input noise to the band width (α) of the smoothing filter.

The justification for using an Edgeworth series must be of a heuristic nature. Cramér[27] has shown that the series is convergent under certain strict conditions and asymptotic for less stringent conditions. Because of the similarity between our expression and Edgeworth series for the sum of N independent quantities, it is believed that the series is asymptotic in the ratio $\beta=\omega_F/\alpha$ (for fixed M) although a proof of this is undoubtedly difficult in our example. In any case, the Edgeworth series with these coefficients gives the correct first four moments of the distribution of M. In practice, the ratio β is very large in order to obtain good smoothing, and the distribution is therefore very nearly normal. Obviously, the same type of argument can be offered for other linear filters, although the evaluation of the higher moments is even more difficult. Similar remarks apply for the distribution of $M(T)=\frac{1}{T}\int_0^T y^2(t)dt$.

APPENDIX B. REMARKS ON OPTIMUM LINEAR FILTERS FOR FINITE-TIME INTEGRATION

As has been seen from Eqs. (2.17) and (2.19), the relative mean square error in the measurement of a wave after it has been smoothed for a finite time T by a low-pass filter, is represented by

$$L\equiv\frac{\langle M(T)^2\rangle-\langle M(T)\rangle^2}{\langle M(T)\rangle^2} =\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_Z(|u-v|)h(u,T)h(v,T)dudv}{\langle Z\rangle^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(u,T)h(v,T)dudv} \qquad (B.1)$$

where ϕ_Z is the autocorrelation function of the input wave to the integrating network, after the steady component $\langle Z\rangle^2$ has been removed, i.e., $\phi_Z(|u-v|)=R_Z(u-v)-\langle Z\rangle^2$ (see Eq. (2.19)). As explained earlier in the text, $h(u,T)$ is the weighting function of the smoothing filter, whose explicit optimum form we wish, such that L above is *minimized*.

To determine the optimum $h(u,T)$, the following variational principle is invoked

$$\delta L=0, \text{ subject to } \int_{-\infty}^{\infty}h(u,T)du=\lambda(>0), \text{ an arbitrary constant.} \qquad (B.2)$$

[25] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946), p. 213.

[26] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, 1946), p. 228.

[27] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, 1946), p. 223.

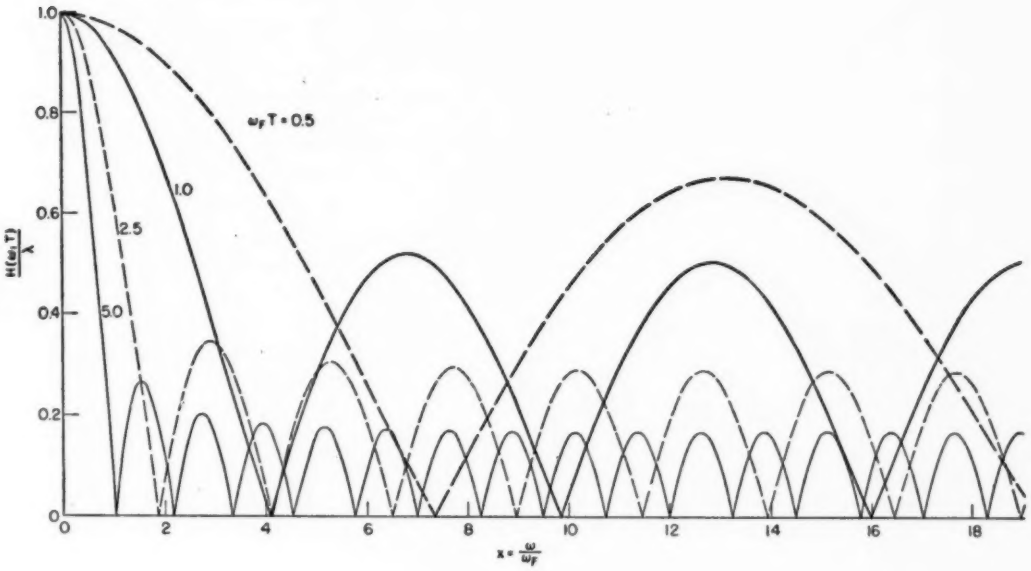


FIG. 3. The frequency response of the optimum filter for representative values of $\omega_F T$.

Because the kernel ϕ_Z is symmetric, one easily finds on applying (B.2) to (B.1) that

$$\delta L = \frac{2}{\langle Z\rangle^2}\int_{-\infty}^{\infty}\delta h(v,T)dv\int_{-\infty}^{\infty}(\phi_Z(|u-v|)-\langle Z\rangle^2 L)h(u)du = 0$$

or

$$\int_{-\infty}^{\infty}\phi_Z(|u-v|)h(u,T)du = L\langle Z\rangle^2\int_{-\infty}^{\infty}h(u,T)du = \lambda\langle Z\rangle^2 L \equiv \Gamma(>0), \text{ a constant.} \quad \text{(B.3)}$$

Equation (B.3) is the integral equation which specifies $h(u, T)$, the optimum filter's time response. Before a solution of (B.3) can be attempted, we must know $\phi_Z(|u-v|)$, the kernel, observing that the complete statement of (B.3) is

$$\int_{-\infty}^{\infty}\phi_Z(|u-v|)h(u,T)du = \begin{cases}\beta_1(v), & T<v\\ \Gamma, & 0<v<T,\\ \beta_2(v), & 0>v\end{cases} \quad \text{(B.4)}$$

where $\beta_1(v)$ and $\beta_2(v)$ are as yet undetermined functions of v.

As a specific example, let us examine from this point of view the problem considered in Part III: The measurement of the power in the normal random noise wave $y(t)$ at the output of a low-pass RC filter with $1/RC=\omega_F$. The autocorrelation function of the fluctuating part of this wave is

$$\phi_Z(|u-v|) = R_Z(u-v) - \langle Z\rangle^2 = 2\langle Z\rangle^2 e^{-2\omega_F|u-v|},$$

(see Eq. (3.1)), where $2\langle Z\rangle^2$ is the mean intensity of the noise continuum.‖ Because the kernel is continuous for all (real) values of u and v, it is evident that $\beta_1(v)=a_1e^{-2\omega_F v}$ for $v>T$ and $\beta_2(v)=a_2e^{2\omega_F v}$, when $v<0$. The arbitrary constants a_1, a_2 are easily determined from the continuous nature of the kernel at $v=0$ and $v=T$ to be $a_1=\Gamma e^{2\omega_F T}$ and $a_2=\Gamma$. The integral equation (B.4) now becomes specifically

$$2\langle Z\rangle^2\int_{-\infty}^{\infty}e^{-2\omega_F|u-v|}h(u,T)du = \begin{cases}\Gamma e^{-2\omega_F(v-T)}, & T<v\\ \Gamma, & 0<v<T.\\ \Gamma e^{+2\omega_F v}, & v<0\end{cases} \quad \text{(B.5)}$$

The solution follows directly upon successive Fourier transformations of both members of (B.5), with the help of the fact that

$$h(u,T)=h(u),\quad 0\le u\le T;\quad h(u,T)=0, \text{ elsewhere,} \quad \text{(B.6)}$$

by the definition of $h(u, T)$, (Eq. (2.14)). One finds, finally, on multiplying both sides of (B.6) by $e^{-i\omega v}$ and integrating, that

$$H(\omega,T)=\int_{-\infty}^{\infty}h(u,T)e^{-i\omega u}du = \frac{\Gamma}{4\langle Z\rangle^2}\{1+e^{-i\omega T}+2\omega_F(1-e^{-i\omega T})/i\omega\}. \quad \text{(B.7)}$$

The inversion of (B.7) then yields for the desired optimum filter¶

$$h(u,T)=\frac{\omega_F\Gamma}{\langle Z\rangle^2}+\frac{\Gamma}{2\langle Z\rangle^2}\{\delta(u-0)+\delta(u-T)\},\quad 0\le u\le T;\ =0, \text{ elsewhere.} \quad \text{(B.8)}$$

There remains the evaluation of Γ and L for this optimum case: Γ follows at once from the condition $\int_0^T h(u)du=\lambda$, and L is obtained in a straightforward way from (B.3). Summarizing these results, we write

$$L=\frac{\sigma_M^2}{\langle M\rangle^2}=\frac{2}{1+\omega_F T},\quad \Gamma=\lambda\langle Z\rangle^2(1+\omega_F T)^{-1},$$
$$h(u,T)=\frac{\lambda\omega_F}{1+\omega_F T}+\frac{\lambda}{1+\omega_F T}\{\delta(u-0)+\delta(u-T)\}. \quad \text{(B.9)}$$

The frequency response obtained by rewriting (B.7) is found to be

$$H(\omega,T)=\frac{\lambda\omega_F T}{1+\omega_F T}\left[\sin\frac{\omega T}{2}\Big/\frac{\omega T}{2}+\cos\frac{\omega T}{2}\Big/\omega_F T\right]e^{-(i\omega T)/2}. \quad \text{(B.10)}$$

Figure 3 shows the frequency characteristic for some typical values of $\omega_F T$.

Since we do not know the phase of the random wave relative to the instant $t=0$ at which integration is begun, our solution that the best weighting of the input is a uniform one within the interval $(0, T)$ is not surprising. Physically, we expect under these circumstances that all points in the interval are to be given the same emphasis, there being no *a priori* reason for preferring one set of times over any other in the range $(0, T)$.

The delta-functions at the beginning and end of the smoothing period are somewhat more difficult to account for physically: A possible explanation is that they are needed for this "equal weighting." At $t=0$ the first *observed* value is completely uncorrelated with any other value of the wave (which is assumed to be stationary and present for all time). From the filter's point of view this point then belongs to a *purely random process*, whereas the actual random wave observed in $(0<t<T)$ [and existing for all t, but unobserved outside this interval and hence of no concern

‖ This expression ($\phi\sim e^{-\alpha|v-u|}$) is known as *Picard's kernel* [see Lalesco, *Introduction à la Théorie des Equations Intégrales*, Hermann et Cie (1912)], and is one for which our integral equation (B.4) can be solved. (We are indebted to Professor Mark Kac for this reference, and to Dr. James Storer, whose solution we follow in part here.)

¶ The delta-functions are assumed here to be symmetrical, e.g., $\int_0^T\delta(u-0)du=\frac{1}{2}$, etc.

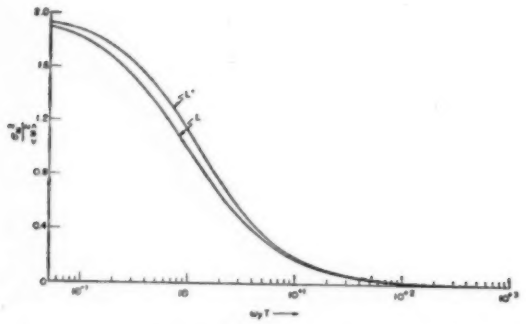


FIG. 4. A comparison of the relative mean square errors obtained from an ideal integrator, L', and the optimum filter, L.

to us] must belong to a Markoff or higher-order process, since the autocorrelation function ϕ_Z is finite. The delta-function "weighting" at $t=0$, as compared with the uniform, finite weighting within $(0, T)$, is needed to compensate for the filter's "interpretation" of the initial value. Similarly, at $t=T$, the filter interprets the wave as being more *highly* correlated than it actually is, and a corresponding compensation is required and we interpret the delta-function here as reducing the influence of this final point.

The relative importance of these end points can be estimated in the case of noise through an *RC* filter by comparing the effects of the optimum and ideal filters** on the relative rms error. Here we have specifically

$$h(u, T)=\lambda/T, \quad (0<u<T); \quad =0, \quad \text{elsewhere.} \tag{B.11}$$

The relative error is then found from (2.3) and (2.6) to be

$$L'=\frac{\langle M(T)^2\rangle-\langle M(T)\rangle^2}{\langle M(T)\rangle^2}=\frac{4\langle Z\rangle^2}{T\langle Z\rangle^2}\int_0^T\left(1-\frac{x}{T}\right)e^{-2\omega_F x}dx$$
$$=\frac{1}{\omega_F{}^2T^2}\{e^{-2\omega_F T}+2\omega_F T-1\}, \tag{B.12}$$

and this is to be compared with $L=[2/(1+\omega_F T)]$, Eq. (B.9). Figure 4 shows L and L' as a function of $\omega_F T$; only when $\omega_F T$ is $0(10^0)$ is there a noticeable difference between the two, and even this difference is quite small, since

$$\lim_{\omega_F T\to 0} L\to 2(1-\omega_F T), \quad \text{and} \quad \lim_{\omega_F T\to 0} L'\to 2(1-\tfrac{2}{3}\omega_F T). \tag{B.13}$$

On the other hand, when $\omega_F T\to\infty$, both L and L' approach zero (as expected) at the same rate, *viz.*: $0(2/\omega_F T)$, and even for smoothing intervals (T) but a few multiples of the time constant $(\omega_F{}^{-1})$, the ideal filter (B.11) gives results essentially identical with those obtained by the true optimum filter (B.9).

Remarks on the optimum filter for a general kernel: Even though the precise form of the kernel in the defining equation (B.4) may not be known, we can prove that *for integration times long* (by a factor 10 at least) *compared to the correlation time*†† T_c *of the random wave in question, the optimum smoothing filter is always the one whose weighting function is a* (nonvanishing) *constant in the integration period* $(0, T)$, in other words, the "ideal" filter (B.11).

From the fact that all filters must give the same result $(L=0)$ in the limit $[\alpha T\to\infty, \alpha T_c\to 0$, where $\alpha=$band width (effective time constant of filter)$^{-1}]$, without any loss of generality we can write with the help of (B.9) for the weighting function of the general optimum filter,

$$h(u, T)=\Gamma'/2+\Psi(u, T), \tag{B.14}$$

where the "end effect," $\Psi(u, T)$, depends of course on the specific form of the kernel ϕ_Z. Let

$$\Phi(v)=\int_v^\infty \phi_Z(u)du, \quad \text{where} \quad \int_0^\infty \phi_Z(u)du=\Phi(0)=\kappa, \tag{B.15}$$

in which κ is for the moment an arbitrary constant. Then from (B.4) we find for the interval $(0\le v\le T)$ that

$$\int_0^T \phi_Z(|u-v|)\left[\frac{\Gamma'}{2}+\Psi(u, T)\right]du=\Gamma, \tag{B.16a}$$

or, on setting $u-v=u'$,

$$\Gamma=\frac{\Gamma'}{2}\left\{\int_0^{T-v}\phi_Z(u')du'+\int_0^v \phi_Z(u')du'\right\}$$
$$+\int_{-v}^{T-v}\phi_Z(u')\Psi(u'+v; T)du', \tag{B.16b}$$

which reduces finally to

$$\int_0^T \phi_Z(|u-v|)\Psi(u, T)du$$
$$=\frac{\Gamma'}{2}\{\Phi(v)+\Phi(T-v)\}+(\Gamma-\Gamma'\kappa) \tag{B.17a}$$
$$=\frac{\Gamma'}{2}\{\Phi(v)+\Phi(T-v)\}, \tag{B.17b}$$

when the arbitrary constant κ is set equal to Γ/Γ' (see (B.10) and (B.14)). We have accordingly transformed the original integral equation for $h(u)$ into one which defines the end effect Ψ. Setting $v'=T-v$ and $u'=T-u$ alternately in (B.17b), we find after a little manipulation that

$$\int_0^T \phi_Z(|u-v|)\Psi(T-u)du=\int_0^T \phi_Z(|u-v|)\Psi(u)du$$
$$=\left(\frac{\Gamma'}{2}\right)[\Phi(v)+\Phi(T-v)], \quad (0\le v\le T). \tag{B.18}$$

If now the smoothing time T is required to be *long* compared to the correlation time T_c, i.e., $T\gg T_c$, then for all v such that $T_c\ll v\ll T-T_c$, $\Phi(v)$ and $\Phi(T-v)$ are negligible, since ϕ_Z (and therefore Ψ by (B.18)) falls off rapidly for times longer than T_c. Only for values of $v=0(T_c)=0(T-T_c)$ is the end-effect of the same order of magnitude as the constant part of the response $\Gamma'/2$; for the vast majority of the interval $(0, T)$, then, the weighting function is a constant. The cumulative effects of the "end-correction" Ψ are always small, if ϕ_Z is continuous, and vanish rapidly relative to the contribution of the constant part of the weighting function when T increases. The *RC* kernel discussed above offers a specific example: $\Psi(u, T)$ here is $\lambda/(1+\omega_F T)[\delta(u-0)]$.

** By ideal filter is here meant one which smooths according to the weighting function λ/T in the interval $(0, T)$, and zero outside this range.

†† The correlation time (for entirely random waves, containing no periodicities or dc component) is here defined as the time required for the correlation function to fall to one-half its (maximum) value at $t=0$.

The Elastic Constants of Nickel

J. R. NEIGHBOURS, F. W. BRATTEN, AND CHARLES S. SMITH
Case Institute of Technology, Cleveland, Ohio

(Received April 12, 1951)

The elastic constants of nickel have been determined by the pulsed ultrasonic method. Wave velocity measurements on four nickel single crystals of general orientation which were magnetically saturated were combined by the approximation method leading to values of the elastic constants $C_{11}=2.53$, $C_{12}=1.52$, $C_{44}=1.24$, all in units of 10^{12} dyne cm^{-2}.

INTRODUCTION

OTHER investigators have previously measured the elastic constants of nickel. Two of these measurements[1,2] were made by combining Young's modulus as a function of crystal direction with the measured value of the cubic compressibility of polycrystalline nickel. Other determinations[3,4] were made by measurement of the three acoustic wave velocities (one longitudinal and two transverse waves) of nickel single crystals cut with their axes in the [110] direction. In this special direction the wave velocities are all different, and furthermore, any one wave may be excited singly by an appropriately oriented transducer.

Until recently most acoustic measurements on cubic crystals have been made on single crystals cut in special directions. However, single crystals with no special orientation were used by Arenberg[5] to determine the elastic constants of silver chloride. The present authors[6] have shown how this method may be extended to higher orders of approximation.

The present paper shows how the approximation method may be applied to combine in a simple manner the wave velocity measurements of several single crystals of different orientation. The elastic constants of nickel have been determined using this method and are compared with the previously determined values in Table I.

EXPERIMENTAL PROCEDURE

Nickel single crystals were prepared by slowly cooling a hydrogen atmosphere furnace containing the nickel charge through the melting point of nickel, the nickel charge being placed in a region of high temperature gradient. A high temperature gradient was obtained by supporting the charged crucible (Norton Company extraction thimble No. 20408, RA-84) on the lip of a $\frac{3}{8}$-in. countersunk hole drilled in a lavite disk. The crucible bottom was allowed to radiate through this hole, thus ensuring an axial heat flow in the charge.

The furnace is a Lindberg CF-1 glowbar heated combustion tube furnace which was mounted so that the combustion tube is vertical. Cooling the furnace was accomplished by slowly decreasing the power supplied. A reduction of temperature from 1600°C to 1450°C in two hours was found to give good results.

The ingots obtained are about $\frac{5}{8}$ in. in diameter and 2 in. long with a weight of approximately 80 grams. Usually there were several single crystals in each ingot, but on two occasions the entire ingot was one single crystal.

A spectrographic analysis of the ingots from which the single crystal cylinders were obtained showed no evidence of impurities except for aluminum, which was estimated to be present in amount less than one-tenth percent. This trace of aluminum is not present in the original nickel charge.[7] Presumably the aluminum was picked up from the Alundum crucible.

The densities as measured by the method of hydrostatic weighing agree with the x-ray value within the experimental error. The lattice constant determined for each crystal is in good agreement with the accepted

TABLE I. Elastic constants of nickel and derived quantities. All measurements were made on magnetically saturated nickel.

Investigator	C_{11} (dyne cm^{-2})	C_{44} (dyne cm^{-2})	$C=C_{11}-C_{12}-2C_{44}$ (dyne cm^{-2})	C_{12} (dyne cm^{-2})	$K=\frac{3}{C_{11}+2C_{12}}$ Compressibility (cm^2 dyne^{-1})	$\frac{2C_{44}}{C_{11}-C_{12}}$ Anisotropy factor	$\frac{C_{11}-C_{12}}{2}$ (dyne cm^{-2})
Bell Telephone Laboratories	2.53×10^{12}	1.22×10^{12}	-1.50×10^{12}	1.58×10^{12}	0.528×10^{-12}	2.57	0.475×10^{12}
Bell Telephone Laboratories	2.524	1.23	−1.47	1.538	0.536	2.49	0.493
Bell Telephone Laboratories	2.523	1.23	−1.50	1.566	0.531	2.57	0.479
This paper	2.52_8	1.23_8	-1.47_1	1.52	0.538 0.535[a]	2.46	0.502

[a] Experimental compressibility.

[1] K. Honda and Y. Shirakawa, Sci. Rep. Res. Inst. Tohoku Univ. **1**, 9 (1949).
[2] M. Yamamoto, Phys. Rev. **77**, 566 (1950).
[3] Bozorth, Mason, and McSkimin, Phys. Rev. **75**, 1954 (1949).
[4] Bozorth, Mason, and McSkimin, "Frequency dependence of elastic constants and losses in nickel," Bell System Tech. J. (October, 1951). The authors are grateful for the opportunity of seeing this paper in manuscript form prior to its publication.
[5] D. Arenberg, J. Appl. Phys. **21**, 941 (1950).
[6] J. R. Neighbours and Charles S. Smith, J. Appl. Phys. **21**, 1338 (1950).
[7] Supplied by Mr. E. M. Wise of the International Nickel Company.

TABLE II. Measured lattice constants and densities of four nickel single crystals.

Specimen	Lattice constant (A)	Density (g cm^{-3})
A	3.5240	8.88
B	3.5238	8.86
C	3.5237	8.89
D	3.5235	8.87
Accepted value	3.5240	8.90

value from which the x-ray density is computed. The authors feel that this is sufficient reason for using the x-ray density (8.90 g cm^{-3}) to compute values of ρV^2 which were subsequently needed. Table II shows the measured densities and lattice constants for four crystals.

After removal from the furnace, the ingots were electropolished and electroetched in nitric acid until the crystal boundaries, if present, were visible. The ingot was then mounted in a collet chuck and a single crystal cylinder was sawed from the ingot using a thin water-cooled abrasive wheel. The end surfaces of the crystal cylinder were then carefully surface ground on a standard machine tool until flat and parallel. The resulting crystal cylinders were about 8 mm long. Acoustic measurements were then made on the crystals with no further attempts to remove the cold-worked layer produced by surface grinding. After the acoustic measurements had been completed, one of the flat surfaces of the cylinder was electropolished to remove the cold-worked layer.

The orientation of the surface normal with respect to the crystallographic axes was then found by the back reflection x-ray method. About ten Laue spots were indexed and measured, and the method of least squares was used to find the direction cosines of the surface normal. The least squares method is quite easy to apply in the case where the direction cosines are regarded as the variables since the coefficients in the observation equations are simply the Miller indices of the reflections. A criterion for these direction cosines is that the sum of their squares should equal one. The greatest error of closure found was 0.001.

The longitudinal and transverse acoustic wave velocities were measured by the pulsed ultrasonic method. The electronic apparatus is similar to that used by Lazarus[8] with the exception that separate quartz crystal transducers are used for sending and receiving. An appropriately cut quartz crystal was attached to each face of the crystal cylinder using phenyl salicylate as a cement. The transmitting quartz was then pulsed to send a short group of 10-mc acoustic waves through the specimen. Because of the poor acoustic impedance match between quartz and nickel, a wave in the nickel crystal was reflected back and forth giving a series of pulses in time. Measurement of the transit time T between echoes and the length L of the crystal cylinder then enables one to calculate the apparent wave velocity, $V_A = 2L/T$.

In general, two transverse waves with different velocities were excited in the nickel crystal by a Y-cut quartz crystal. However, by rotating the plane of polarization one of these could be made very small. The velocity of the predominant wave was then measured. By further rotation, the dominant and minor roles were interchanged, and the velocity of the other wave was then measured.

Attenuation in the nickel crystals is so large that first attempts to measure wave velocities were inconclusive. However, when the nickel crystal was placed in a transverse saturation magnetic field, the attenuation was diminished so that upwards of twenty echoes could be seen on the measuring oscilloscope (Dumont 256D) screen. All velocity measurements were made in a transverse magnetic field conveniently supplied by an Alnico magnetron magnet. The cylindrical face of the nickel crystal was magnetically held in contact with one pole face of the magnetron magnet. The field in the gap was determined to be 5 kilo-oersted which value is roughly double that of the demagnetizing field for the specimens. Qualitative experiments conducted here in a variable magnetic field (and quantitatively carried out elsewhere[4]) showed that the damping "saturated" at very low field strengths. It was therefore assumed that the nickel crystals were magnetically saturated and that the elastic constants to be derived were characteristic of such material.

In these qualitative experiments involving the magnetic field it was not possible to observe the ΔE effect[4] (appearing as about a 1.5 percent effect in velocity) because of the low precision of our no field velocity. The second-order ΔE effect[9] (about 0.1 percent in velocity) was not observed either, but its magnitude is on the edge of the precision of the present work.

Other investigators[10] in measuring transit times, state that they have measured to the front of the echo, thus assuring that the acoustic path involved was contained entirely within the specimen. Lazarus,[8] however, men-

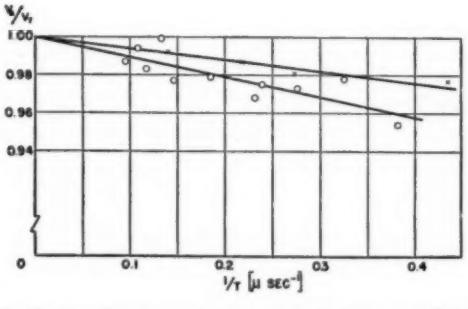


FIG. 1. Ratio of apparent to true velocity as a function of reciprocal transit time for magnetically saturated polycrystalline nickel. The upper curve (crosses) is for longitudinal waves and the lower (circles) is for transverse waves.

[8] David Lazarus, Phys. Rev. 76, 545 (1949).

[9] W. P. Mason, Phys. Rev. 82, 715 (1951).

[10] H. B. Huntington, Phys. Rev. 72, 321 (1947).

tions finding a small transit time error which is ascribed to the travel time in the cement and transducer. This laboratory has followed the procedure of measuring to the front of the pulse and has also found a transit time error which in the present case is particularly important because of the short length of the specimens. Extensive experimentation has not led to a clear understanding of the origin of the error and so the policy of experimentally determining it has been adopted.

The transit time error was determined by measuring transit times for both longitudinal and transverse waves in magnetically saturated polycrystalline nickel for several different specimen lengths. Except for the material being polycrystalline the experimental conditions were the same as those used for the nickel crystals. The observed transit times were plotted against the measured lengths and a least squares line was fitted to the data. The intercept of this line is the assumed transit time error ΔT, and the reciprocal of the slope is proportional to the true velocity V_T. The values of ΔT are found to be $\Delta T=0.061\pm0.005$ μsec for longitudinal waves and $\Delta T=0.106\pm0.024$ for transverse waves.

The true velocity is given by $V_T=2L/(T-\Delta T)$, and we note that the ratio of apparent velocity to true velocity is $V_A/V_T=1-\Delta T/T$. This relation, shown in Fig. 1 for the measurements on polycrystalline nickel, is more convenient for graphically displaying the systematic effect than the simpler relation used for the least squares determination.

The precision of the determination of an apparent velocity is estimated to be of the order of 0.2 percent. The systematic transit time error amounts to about 2.5 percent in the case of both the longitudinal and transverse velocities. Applying the probable errors quoted above leads one to the conclusion that the "true" velocities are uncertain because of the systematic error alone by ±0.3 percent in the longitudinal case and by ±0.6 percent in the transverse case. Combining these figures with the estimated precision in the usual way and converting the probable errors to those in ρV^2 leads to the expected probable errors of ±0.6 percent in longitudinal ρV^2 and ±1.2 percent in transverse ρV^2.

Calculation of the true wave velocities in nickel crystals was made assuming that the polycrystalline values of ΔT applied. Table III is a listing of orientation function, Γ, and the corrected values of ρV^2 for the four single crystals.

TABLE III. Measured orientation functions and values of ρV^2. The values of ρV^2 are computed using the x-ray density. The subscript 1 refers to longitudinal waves, subscripts 2 and 3 refer to transverse waves.

Specimen	Orientation function, Γ	Longitudinal ρV_1^2 (dyne cm^{-2})	Transverse ρV_2^2 (dyne cm^{-2})	Transverse ρV_3^2 (dyne cm^{-2})
A	0.3253	3.482×10^{12}	0.655×10^{12}	0.843×10^{12}
B	0.1096	2.942	1.185	0.888
C	0.2138	3.225	1.050	0.754
D	0.3044	3.431	0.943	0.617

CALCULATION OF ELASTIC CONSTANTS

One faces a peculiar problem in determining the elastic constants of a cubic crystal from measurements of wave velocities in an arbitrary general direction. The values of ρV^2 (where ρ is the density) are the roots of the secular equation for the given direction. Various combinations of the three elastic constants make up the coefficients of the secular equation, and the problem is to determine a set of elastic constants such that the solution of the secular equation then gives the experimental values of ρV^2.

The authors have recently given such a method[6] applicable to measurements on one single crystal. If two or more crystals are available, the labor may be considerably reduced by the following modifications. From this previous paper, Eq. (6), the roots of the secular equation as given by the perturbation method are

$$\left\{\begin{aligned}
\rho V_1^2 &= C_{11}'+\left\{\left[\frac{C_{15}'^2}{C_{11}'-C_{55}'}+\frac{C_{16}'^2}{C_{11}'-C_{66}'}\right]+\left[\frac{2C_{15}'C_{16}'C_{56}'}{(C_{11}'-C_{55}')(C_{11}'-C_{66}')}\right]+\cdots\right\}\\
\rho V_2^2 &= C_{66}'+\left\{\left[\frac{C_{16}'^2}{C_{66}'-C_{11}'}+\frac{C_{56}'^2}{C_{66}'-C_{55}'}\right]+\left[\frac{2C_{15}'C_{16}'C_{56}'}{(C_{66}'-C_{11}')(C_{66}'-C_{55}')}\right]+\cdots\right\}\\
\rho V_3^2 &= C_{55}'+\left\{\left[\frac{C_{56}'^2}{C_{55}'-C_{66}'}+\frac{C_{15}'^2}{C_{55}'-C_{11}'}\right]+\left[\frac{2C_{15}'C_{16}'C_{56}'}{(C_{55}'-C_{66}')(C_{55}'-C_{11}')}\right]+\cdots\right\},
\end{aligned}\right. \tag{1}$$

and the primed elastic constants as given by Eq. (4) are

$$\left.\begin{aligned}
C_{11}' &= C_{11}-2[l_1^2m_1^2+l_1^2n_1^2+m_1^2n_1^2]C\\
C_{16}' &= l_1m_1[m_1^2-l_1^2][l_1^2+m_1^2]^{-1}C\\
C_{66}' &= C_{44}+2l_1^2m_1^2[l_1^2+m_1^2]^{-1}C\\
C_{15}' &= -n_1[l_1^4+m_1^4-n_1^2(l_1^2+m_1^2)][l_1^2+m_1^2]^{-1}C\\
C_{55}' &= C_{44}+2n_1^2[l_1^4+l_1^2m_1^2+m_1^4][l_1^2+m_1^2]^{-1}C\\
C_{56}' &= l_1m_1n_1[l_1^2-m_1^2][l_1^2+m_1^2]^{-1}C\\
C &= [C_{11}-C_{12}-2C_{44}].
\end{aligned}\right\} \tag{2}$$

For simplicity we may rewrite Eqs. (2)

$$\left.\begin{aligned}
C_{11}' &= C_{11}-2\Gamma C\\
C_{66}' &= C_{44}+a_1C\\
C_{55}' &= C_{44}+a_2C\\
C_{15}' &= -b_1C\\
C_{16}' &= b_2C\\
C_{56}' &= b_3C.
\end{aligned}\right\} \tag{3}$$

It may be shown that $a_1+a_2=2\Gamma$.

A rough calculation quickly shows that the third-order perturbation term for nickel is negligible. Adding the last two of Eqs. (1) and neglecting perturbation terms of higher order than second we may write

$$\left.\begin{aligned}\rho V_1^2 &= C_{11}-2\Gamma C \\ &\quad +\left\{\frac{C_{15}'^2}{C_{11}'-C_{55}'}+\frac{C_{16}'^2}{C_{11}'-C_{66}'}\right\} \\ \rho V_2^2+\rho V_3^2 &= 2C_{44}+2\Gamma C \\ &\quad -\left\{\frac{C_{15}'^2}{C_{11}'-C_{55}'}+\frac{C_{16}'^2}{C_{11}'-C_{66}'}\right\}.\end{aligned}\right\} \quad (4)$$

We may neglect the perturbation corrections where they appear in the denominator of the perturbation corrections themselves in Eqs. (4). That is, we may use the approximate equalities,

$$\left.\begin{aligned}C_{11}'-C_{55}' &\sim \rho V_1^2-\rho V_3^2+\frac{b_3^2C^2}{\rho V_3^2-\rho V_2^2} \\ &\quad -\frac{b_2^2C^2}{\rho V_1^2-\rho V_2^2}-\frac{2b_1^2C^2}{\rho V_1^2-\rho V_3^2} \\ C_{11}'-C_{66}' &\sim \rho V_1^2-\rho V_2^2+\frac{b_3^2C^2}{\rho V_2^2-\rho V_3^2} \\ &\quad -\frac{b_1^2C^2}{\rho V_1^2-\rho V_3^2}-\frac{2b_2^2C^2}{\rho V_1^2-\rho V_2^2},\end{aligned}\right\} \quad (5)$$

which are derived by combining Eqs. (1) and (3).

Then substituting Eqs. (5) into the denominators of Eqs. (4) and expanding by the binomial theorem, we arrive at the formulas used to compute the elastic constants.

$$\left.\begin{aligned}F(\Gamma) &= \rho V_1^2-\{AC^2+BC^4\}=C_{11}-2\Gamma C \\ G(\Gamma) &= \rho V_2^2+\rho V_3^2+\{AC^2+BC^4\}=2C_{44}+2\Gamma C \\ A &= \frac{b_1^2}{\rho V_1^2-\rho V_3^2}+\frac{b_2^2}{\rho V_1^2-\rho V_2^2} \\ B &= \frac{b_1^2}{(\rho V_1^2-\rho V_3^2)^2}\left[\frac{b_3^2}{\rho V_3^2-\rho V_2^2}\right. \\ &\quad \left.-\frac{b_2^2}{\rho V_1^2-\rho V_2^2}-\frac{2b_1^2}{\rho V_1^2-\rho V_3^2}\right] \\ &\quad +\frac{b_2^2}{(\rho V_1^2-\rho V_2^2)^2}\left[\frac{b_3^2}{\rho V_2^2-\rho V_3^2}\right. \\ &\quad \left.-\frac{b_1^2}{\rho V_1^2-\rho V_3^2}-\frac{2b_2^2}{\rho V_1^2-\rho V_2^2}\right].\end{aligned}\right\} \quad (6)$$

The constants A and B are computed for each crystal. The measured values of ρV_1^2 and $(\rho V_2^2+\rho V_3^2)$ are then plotted against Γ. These are the first approximations for $F(\Gamma)$ and $G(\Gamma)$. The slopes of these lines give a trial value for C. This value of C is then used to compute the correction term $\{AC^2+BC^4\}$ which then allows one to calculate a second approximation for $F(\Gamma)$ and $G(\Gamma)$. These new values of $F(\Gamma)$ and $G(\Gamma)$ are then plotted against Γ. The slopes of these lines then give a corrected value for C.

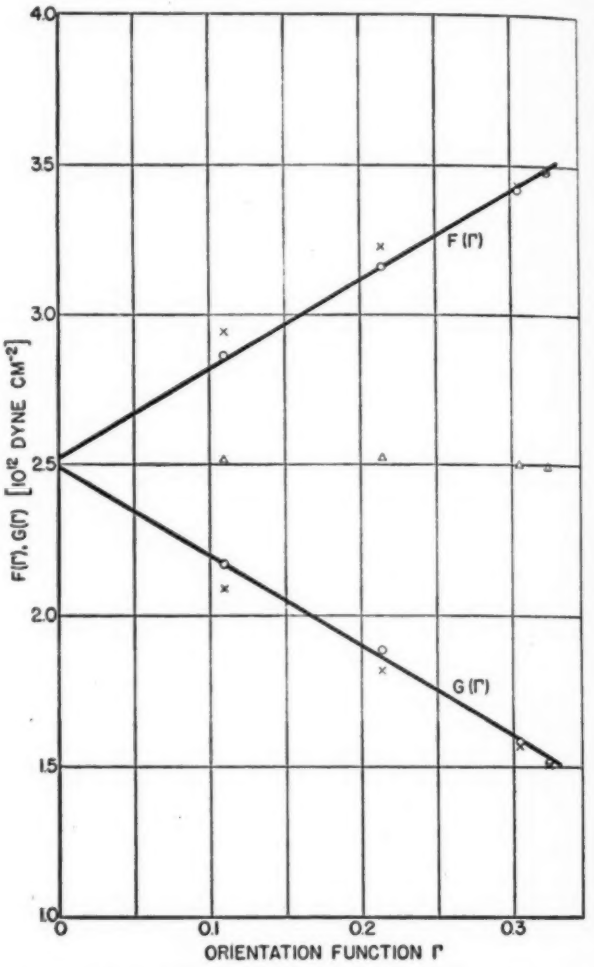


FIG. 2. The functions $F(\Gamma)$ and $G(\Gamma)$ *vs* Γ: The Δ points are one-half the sum of the three values of ρV^2. The upper $\times$ points are measured values of ρV_1^2 (longitudinal waves), the first approximation to $F(\Gamma)$. The nearby $\bigcirc$ points are final corrected values of $F(\Gamma)$ or C_{11}'. The lower $\times$ points are measured values of $(\rho V_2^2+\rho V_3^2)$ (transverse waves), the first approximations to $G(\Gamma)$. The nearby $\bigcirc$ points are final corrected values of $G(\Gamma)$, or $(C_{55}'+C_{66}')$.

This process may be repeated until the values of C obtained approach a self-consistent value. The constants C_{11} and C_{44} may then be determined from the intercepts of the two lines. Usually the line through the $F(\Gamma)$ points will not have the same slope as the line through the $G(\Gamma)$ points. However, since the material has only one constant C, it is proper to impose the condition

that the two slopes be equal. It is convenient to impose this condition, and in fact to carry out the whole iterated computation, by using the least squares method from the beginning. The results have actually been computed in this way, employing eight observation equations to give three normal equations which may be solved for the least squares adjusted values of C_{11}, C_{44}, and C.

Figure 2 is a plot of $F(\Gamma)$ and $G(\Gamma)$ *vs* Γ. The initial plot of ρV_1^2 and $(\rho V_2^2+\rho V_3^2)$ are shown as crosses. The final corrected points $\rho V_1^2-\{AC^2+BC^4\}$ and $(\rho V_2^2+\rho V_3^2)+\{AC^2+BC^4\}$ are shown as circles. The two lines represent the least squares solution.

We note that the sum of the roots of the secular determinant is a constant.

$$F(\Gamma)+G(\Gamma)=\rho V_1^2+\rho V_2^2+\rho V_3^2=C_{11}+2C_{44}. \quad (7)$$

This equation affords a useful check on the internal consistency of the wave velocity determinations on different single crystals. The experimental values of one-half this sum are exhibited as triangles in Fig. 2. The probable error of these experimental values is ±0.3 percent.

RESULTS AND DISCUSSION

The results of the data reduction of the previous section are displayed in Fig. 2, and numerical values of the adiabatic elastic constants at 25°C of magnetically saturated nickel are entered in Table I. There C_{11}, C_{44}, and C may be regarded as directly determined, and the other entries as derived quantities. The internal consistency shown on Fig. 2 justifies the method of data reduction that has been followed and the approximations that have been made in the process. Two things may be noted in connection with Fig. 2. The "perturbation corrections" (the distance between circles and crosses) are not in general a simple monotonic function of Γ as would appear to be the case here. Also there is a slight but definite difference between the longitudinal and transverse points with regard to the slopes of the lines which would pass through each set alone.

Apart from the internal consistency of the present measurements on four crystals is the question of the accuracy of the derived constants. The uncertainty of the values of C_{11}, C_{44}, and C arises almost entirely from the uncertainty of the two transit time corrections which have been made. Starting from the probable errors specified in the experimental section a lengthy but straightforward calculation[11] formally yields for the probable errors: C_{11}, ±0.02; C_{44}, ±0.01; C, ±0.03 all in units of 10^{12} dyne cm^{-2}.

In Table I are quoted for comparison values of C_{11}, C_{44}, and C_{12} recently determined at the Bell Telephone Laboratories. The other entries have been computed by the writers from these values. These constants are also the adiabatic constants of magnetically saturated crystals and so may be directly compared with the present ones. The Bell Telephone Laboratories experiments differ from the present ones in two respects; their three crystals were oriented with their surface normals in the [110] direction, and their velocities were measured in a different manner but at the same frequency (10 mc). The agreement between the determinations from two different laboratories is seen to be good and certainly well within the combined experimental error.

The experimental isothermal compressibility of nickel as determined by Bridgman[12] and modified by Slater[13] is also shown in Table I. The agreement between the experimental compressibility and that computed from the four sets of elastic constants is satisfactory.

The results of Yamamoto[2] and of Honda and Shirakawa[1] on unmagnetized crystals cannot be compared directly with the present results as discussed elsewhere.[4]

The authors wish to thank Dr. L. L. Foldy for his many suggestions and J. W. Burns for his assistance in measuring the wave velocities.

This work was sponsored in part by the Office of Naval Research and in part by the Dayton C. Miller Research Fund of Case Institute of Technology.

[11] A. G. Worthing and J. Geffner, *Treatment of Experimental Data* (John Wiley and Sons, Inc., New York, 1943), p. 229.
[12] P. W. Bridgman, Proc. Am. Acad. Sci. **58**, 165 (1923).
[13] J. C. Slater, Phys. Rev. **57**, 744 (1940).

Effect of Thermal-Velocity Spread on the Noise Figure in Traveling-Wave Tubes*

PHILIP PARZEN
Federal Telecommunication Laboratories, Incorporated, Nutley, New Jersey
(Received July 13, 1951)

Effects of thermal-velocity spread on the noise figure of a traveling-wave tube have been computed by extending a method of Hahn. To accomplish this, it has been necessary to work out a new theory for the diode-drift tube and traveling-wave tube that includes the effect of the thermal-velocity spread at very high frequencies. Some computations are given for conditions met in practice.

CALCULATION of the noise figure of a traveling-wave-tube amplifier may be reduced to the computation of the current- and velocity-modulation fluctuations in three distinct regions (Fig. 1). Region *I* is a diode with a heated cathode, region *II* is a field-free drift region, and region *III* is the traveling-wave-tube amplifier.

A detailed analysis will be given of the fluctuations in each region, which includes the effect of the thermal-velocity spread. Inasmuch as this theory applies rigorously only to an infinite rectilinear electron beam, it will be necessary to apply certain correction factors to account for the finite beam and the surrounding conductors. Approximations for these correction factors may be obtained from the simpler computation for a monovelocity beam.

Of the three regions, the diode region is the most difficult to handle since it is in a dc accelerating field. As is well known, the dc theory of the diode by Langmuir[1] is at best a numerical theory, and certainly not much more can be expected for an ac theory. To attempt to work out an ac theory by Langmuir's method does not seem fruitful for microwave frequencies. However, such theories have been worked out for low frequencies by North *et al.*[2] and Rack.[3]

Rack has extended his theory to higher frequencies by using Llewellyn's[4] theory of diodes in an approximate manner. On the basis of the Rack theory Peterson[5] has computed the fluctuations in the drift tube. This was utilized by Pierce[6] to compute the fluctuations in the traveling-wave amplifier. Subsequently, this theory was extended to include the effect of the space-charge waves discussed by Hahn[7] and Ramo.[8] This method of calculation of the fluctuations in the traveling-wave amplifier will hereafter be designated as the RPHP (Rack-Peterson-Hahn-Pierce) method. In view of the many approximations made, it is not possible to assess the accuracy of this theory except by comparison with experiment. The agreement is very good for large noise figures.

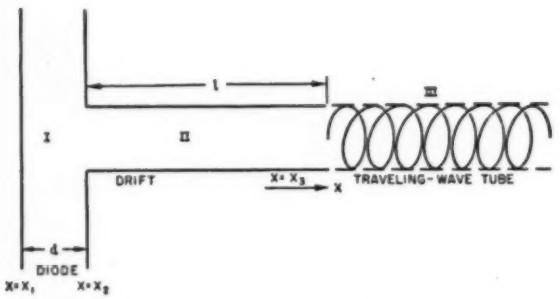


FIG. 1. Nomenclature for noise-figure calculations on a traveling-wave tube.

For noise figures of the order of 10 db or less, however, one cannot expect experimental agreement within 1 db with this theory, unless it be fortuitous. It is expected that the weakest link of the RPHP theory is the diode region. Thus a high frequency theory of diode impedance and fluctuations including the effect of a multivelocity beam would be quite useful.

Another method of attack is to base the ac theory on the dc theory of Hahn.[9] Hahn has shown that it is possible to discuss the effect of the thermal velocities on the dc behavior of diodes in terms of a monovelocity electron beam whose acceleration is determined not only by the applied electric field but also by a hydrostatic-pressure term, which accounts for the multivelocity beam. The ac theory will be worked out as a perturbation of the dc properties determined by the Hahn theory.

1. AC THEORY OF DIODE IMPEDANCE AND FLUCTUATIONS

1.1 General

The diode consists of two infinite parallel planes; the space between is traversed by an infinite rectilinear beam. The diode is connected either to an ac generator or a load Z_L (Figs. 2 and 3).

* This work was sponsored by the Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey.

[1] I. Langmuir, Phys. Rev. **17**, 441 (1921).
[2] North, *et al.*, RCA Rev. **5**, 106 (1940).
[3] A. J. Rack, Bell System Tech. J., **17**, 592 (1938).
[4] F. B. Llewellyn, *Electron Inertia Effects* (Cambridge University Press, London, England, 1941).
[5] L. C. Peterson, Proc. Inst. Radio Engrs. **35**, 1264 (1947).
[6] J. R. Pierce, *Traveling-Wave Tubes* (D. Van Nostrand Company, Inc., New York, 1950), p. 109.
[7] W. C. Hahn, Gen. Elec. Rev. **42**, 258 (1939).
[8] S. Ramo, Proc. Inst. Radio Engrs. **27**, 757 (1939).
[9] W. C. Hahn, Proc. Inst. Radio Engr. **36**, 1115 (1948).

Then, following Hahn, the basic equations are

$$\bar{v}_t+\bar{v}\bar{v}_x=\frac{e}{m}\bar{E}-\frac{k\bar{T}}{mn}\bar{n}_x, \tag{1}$$

$$\bar{J}_x+e\bar{n}_t=0, \tag{2}$$

$$\bar{E}_x=e\bar{n}/\epsilon_0 \quad \text{or} \quad \bar{J}+\epsilon_0\bar{E}_t=I, \tag{3}$$

$$\bar{J}=e\bar{n}\bar{v}. \tag{4}$$

The subscripts t and x denote the partial derivatives with respect to t and x. A bar over the letter denotes total unnormalized quantities, i.e., dc+ac. e and m are the charge and mass of an electron. n, v, J, and E are particle density, mean velocity, convection-current density, and electric field, respectively. ϵ_0 is the dielectric constant of free space. I is the total current density, i.e., convection plus displacement. T is the temperature of the electron gas[10] that constitutes the electron beam. k is Boltzmann's constant. Meter-kilogram-second units are used.

All field quantities are functions of x and t except $\bar{I}$, which is a function of t only and $\bar{T}$ which is assumed to be constant. Thus $\bar{T}=T_0$, which is assumed to equal the cathode termperature. The major justification for this assumption is that the dc potential distribution is approximately the same as derived by Langmuir, who assumes an adiabatic process. So far as the ac potential distribution is concerned, it remains an assumption. It is seen that the equation of motion Eq. (1) has been augmented with a hydrostatic pressure term.

1.2 Small-Signal Approximations

According to the small-signal approximations, one puts

$$\bar{J}=J_0+J\exp j\omega t; \quad J\ll J_0 \tag{5}$$

and similarly for all other field quantities. The subscript zero indicates dc quantities and ω is the applied frequency. It is also assumed that the ac temperature is zero. The ac and dc equations separate out.

Ac equations

$$(j\omega+v_{0x})v+v_0v_x=\frac{e}{m}E-\frac{kT_0}{mn_0}n_x+\frac{kT_0}{mn_0^2}n_{0x}n, \tag{1a}$$

$$J_x+j\omega en=0, \tag{2a}$$

$$J+j\omega\epsilon_0E=I, \tag{3a}$$

$$J=e(n_0v+nv_0). \tag{4a}$$

Dc equations

$$v_0v_{0x}=\frac{e}{m}E_0+\frac{kT_0}{m}\frac{1}{v_0}v_x, \tag{1b}$$

$$E_{0x}=J_0/\epsilon_0v_0, \tag{3b}$$

$$J_0=en_0v_0. \tag{4b}$$

[10] P. Parzen and L. Goldstein, J. Appl. Phys. 22, 398 (1951).

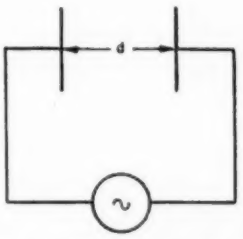


FIG. 2. Diode connected to an ac generator.

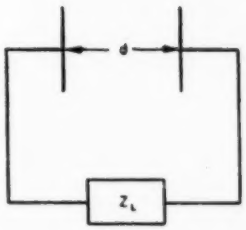


FIG. 3. Diode connected to a load Z_L.

The ac quantities are all functions of x except I which is a constant. The dc quantities are all functions of x except J_0 and T_0 which are constants. It is now possible to solve for all the ac quantities in terms of J and J_x. Thus,

$$v=\frac{1}{J_0}\left(v_0J+\frac{v_0^2}{j\omega}J_x\right), \tag{6}$$

$$n=-\frac{1}{j\omega\epsilon_0}J_x, \tag{7}$$

$$E=\frac{I}{j\omega\epsilon_0}-\frac{J}{j\omega\epsilon_0}, \tag{8}$$

$$\frac{v_0}{j\omega J_0}\left(\frac{kT_0}{m}-v_0^2\right)J_{xx}+\frac{1}{J_0}\left(\frac{1}{j\omega}\frac{kT_0}{m}v_{0x}-\frac{3}{j\omega}v_0^2v_{0x}-2v_0^2\right)J_x+\left(-\frac{e}{m}\frac{1}{j\omega\epsilon_0}-\frac{2v_0}{J_0}v_{0x}-j\omega\frac{v_0}{J_0}\right)J=-\frac{e}{m}\frac{1}{j\omega\epsilon_0}I. \tag{9}$$

These equations can be somewhat simplified by normalizing the units of distance, velocity, and current density. Let

$$y=(x-x_c)/D, \tag{10}$$

where x_c is the position of the virtual cathode, and let the unit of velocity α be such that

$$\frac{e}{m}\frac{1}{\epsilon_0}\frac{D^2J_0}{\alpha^3}=\frac{2}{9}, \tag{11}$$

where D is the unit of distance for normalization and J_0 is the unit of current density. Then, in terms of these normalized units

$$v=v_0J+\frac{v_0^2}{j\theta}J_y, \tag{12}$$

$$v_0(\mu^2-v_0^2)J_{yy}+(\mu^2v_{0y}-3v_0^2v_{0y}-j2\theta v_0^2)J_y+[-(2/9)-2j\theta v_0v_{0y}+\theta^2v_0]J=-(2/9)I, \tag{13}$$

where

$$\theta=\omega D/\alpha \quad \text{and} \quad \mu^2=kT_0/m\alpha^2. \tag{14}$$

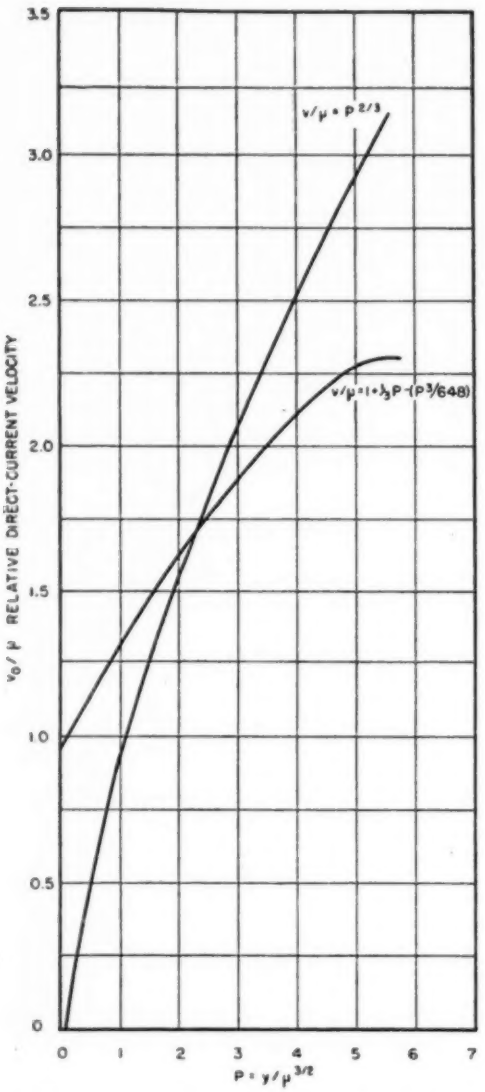


FIG. 4. Merging of Child's law region in the cathode region.

For different regions one may choose different values of D and α subject to Eq. (11). Hereafter, only normalized quantities will be used unless otherwise specified.

1.3 Solution of the dc Equations

It follows immediately from Eq. (1b) that at the virtual cathode, at which $y=0$, $E_0=0$, $v_0=\mu$. In general, v_0 satisfies the following integral equation

$$v_0^2=\mu^2\left(1+\log\frac{v_0^2}{\mu^2}\right)+\frac{4}{9}\int_0^y dy\int_0^y\frac{dy}{v_0}. \tag{15}$$

This equation cannot be solved without lengthy numerical computations, except for the two following special cases: For $|y|\ll 1$

$$v_0=\mu+\frac{1}{3\mu^{1/2}}y+\frac{1}{648\mu^{7/2}}y^2+\cdots, \tag{16}$$

which is obtained by assuming for v_0 a power series expansion around $y=0$. For those values of y for which $(v_0/\mu)^2\gg 1+\log(v_0/\mu)^2$,

$$v_0=y^{\frac{1}{2}}. \tag{17}$$

This will occur for sufficiently large positive values of y, and it corresponds to the region in which Child's law is valid. Inasmuch as there is no simple expression for the region in between, it has been decided to fill in this region by matching Eq. (16) to Eq. (17). It is seen from Fig. 4 that they intersect at $y=y_m=2.4\mu^{\frac{1}{2}}$ at which $v_0=1.76\mu$. Thus the dc solution, of which the ac solution is a perturbation, will be taken as follows:

$$\begin{aligned}&\text{for } y<2.4\mu^{\frac{1}{2}},\ v_0 \text{ is given by Eq. (16)}\\ &\text{for } y>2.4\mu^{\frac{1}{2}},\ v_0 \text{ is given by Eq. (17).}\end{aligned} \tag{18}$$

All these assumptions will be approximately satisfied if the ratio of current density to saturation current density is not too small. The voltage, V, relative to the virtual cathode in the region $y<0$ is

$$V=\frac{4}{9}\frac{V_0}{\mu}\left(\frac{y^2}{2}-\frac{1}{18\mu^{\frac{1}{2}}}y^3+\frac{1}{108\mu^3}y^4+\cdots\right)$$

$$V_0=m\alpha^2/2|e|. \tag{19}$$

Let subscripts 1 and 2 denote conditions at the cathode and anode, respectively.

$$eV_1/kT=\log J_0/J_s, \tag{20}$$

J_s being the saturation current density.

Choosing α such that $4eV_0/9kT=-1$, it follows from Eq. (20) that

$$y_1=-\left(2\log\frac{J_s}{J_0}\right)^{\frac{1}{2}}\left[1-0.12\left(2\log\frac{J_s}{J_0}\right)^{\frac{1}{2}}\right]^{\frac{1}{2}}, \tag{21}$$

$$\alpha=\left(\frac{9}{2}\right)^{\frac{1}{2}}\left(\frac{kT_0}{m}\right)^{\frac{1}{2}},$$

$$D=\left[\frac{\epsilon_0}{eJ_0}\left(\frac{kT_0}{m}\right)^{\frac{1}{2}}kT_0\right]^{\frac{1}{2}}. \tag{22}$$

With this choice of D and α, y is the same as Hahn's s and Langmuir's $\zeta=1.583y$. Also, $\mu=0.61$. Thus given T_0, J_0, and d, one may compute y_1 from Eq. (21), D from Eq. (22), and hence x_c. Then knowing y_2, one computes V_2 from Eqs. (16) and (17), and V_1 from Eq. (20). Thus V_2-V_1 is the required voltage. Hahn has already shown that for $y>0$, V computed from his theory agrees fairly well with that obtained by Langmuir. The agreement in the region $y<0$ between this theory and that of Langmuir is shown in Fig. 5 in which y_1 is plotted against J_s/J_0. It is seen that for $J_s/J_0<3$, the agreement is fairly good. An exact solution of the dc equation is given in the appendix. It will not be used here.

1.4 Solution of the ac Equation

The ac equation is

$$v_0(\mu^2-v_0^2)J_{yy}+(\mu^2 v_{0y}-3v_0^2 v_{0y}-j2\theta v_0^2)J_y +[-(2/9)-2j\theta v_0 v_{0y}+\theta^2 v_0]J=-(2/9)I. \quad (13)$$

It should be noted that this differential equation has a singularity at $y=0$. In view of this, the homogeneous differential equation ($I=0$) possesses only one regular solution at $y=0$. Thus there are only two constants to be determined, I and the arbitrary constant of the regular solution. The required initial conditions will be given later.

An approximate solution of Eq. (13) is obtained as follows. Rewriting Eq. (13) in the form

$$y^2\sum_0 A_i y^i J_{yy}+y\sum_0 B_i y^i J_y+y\sum_0 C_i y^i J=-(2/9)Iy. \quad (23)$$

From Eq. (16)

$$A_0=-\tfrac{2}{3}\mu^{\frac{1}{2}}$$

$$A_1=-\tfrac{1}{3}$$

$$B_0=\mu^{\frac{1}{2}}(-\tfrac{2}{3}-j2\epsilon)$$

$$B_1=-\tfrac{2}{3}(1+j2\epsilon)$$

$$C_0=-2/9[1-(9/2)\epsilon^2+j3\epsilon]$$

$$C_1=-\frac{2}{9}\frac{1}{\mu^{\frac{1}{2}}}\epsilon(-\tfrac{3}{2}\epsilon+j)$$

$$\epsilon=\theta\mu^{\frac{1}{2}}=0.47\omega/\omega_p$$

$$=2\times 10^{-12}\omega\frac{T_0^{\frac{1}{2}}}{|J_0|^{\frac{1}{2}}_{\text{amp/cm}^2}}, \quad (24)$$

where ω_p is the plasma frequency at the virtual cathode. One assumes a power-series solution.

$$J=\sum_0 a_n y^n \quad (25)$$

in the region $y_1<y<2.4\mu^{\frac{1}{2}}$. Equating coefficients of powers of y, one has

$$a_1=-\frac{C_0}{B_0}a_0-\frac{2}{9}\frac{I}{B_0},$$

$$a_2=\frac{a_0}{2(A_0+B_0)}\left[-C_1+(B_1+C_0)\frac{C_0}{B_0}\right] +\frac{B_1+C_0}{2(A_0+B_0)}\frac{2}{9}\frac{I}{B_0}. \quad (26)$$

This expansion will not be carried out further here.

It will be assumed that in view of the smallness of y the first three terms of the series suffice to approximate J in this region. In the region $y>2.4\mu^{\frac{1}{2}}$ a solution is obtained in the following manner. In this region v_0 is given by Eq. (17), which implies physically that the effects of the initial thermal-velocity spread may be neglected. It will be assumed that this is still true for the ac process, since it is a perturbation of the dc process. This is equivalent to putting $\mu=0$ in Eq. (13). The solution of Eq. (13) is now

$$J=\left(\frac{c_1}{Z}+\frac{c_2}{Z^2}\right)\exp(-j3\theta Z)-\frac{2I}{9\theta^2 Z^2}, \quad (27)$$

$$Z^3=y, \quad (28)$$

$$v=-\left(\frac{c_1}{j3\theta}+\frac{2c_2}{j3\theta Z}\right)\exp(-j3\theta Z) -\frac{2I}{9\theta^2 Z^2}\left(Z^2-\frac{2Z}{j3\theta}\right). \quad (29)$$

c_1 and c_2 are computed by matching the value of J_m and J_{my} at $y=y_m=2.4\mu^{\frac{1}{2}}$ of this solution to that derived from Eq. (25). Thus

$$J_m=a_0+a_1 y_m+a_2 y_m^2, \qquad J_{my}=a_1+2a_2 y_m, \quad (30)$$

and

$$c_1=(2Z_m J_m+j3\theta v_m)\exp(j3\theta Z_m)-\frac{2I}{j3\theta}\exp(j3\theta Z_m), \quad (31)$$

$$c_2=(-Z_m^2 J_m-j3\theta Z_m v_m)\exp(j3\theta Z_m) +2I\left(\frac{Z_m}{j3\theta}+\frac{1}{9\theta^2}\right)\exp(j3\theta Z_m),$$

$$v_m=Z_m^2\left(J_m+\frac{J_{mz}}{j3\theta}\right). \quad (32)$$

The transit angle B_{m2} from matching point to anode is equal to

$$3\theta(Z_2-Z_m)=3.9\epsilon(Z_2-Z_m). \quad (33)$$

Thus the solution of the ac problem has been reduced to the determination of a_0 and I. It is here that the problems of ac impedance (Fig. 2) and fluctuations (Fig. 3) diverge. It is to be noted from Eq. (24) that the frequency appears only in the quantity ϵ. For a given set of dc conditions, the ac behavior of a diode† is determined solely by the ratio of the applied fre-

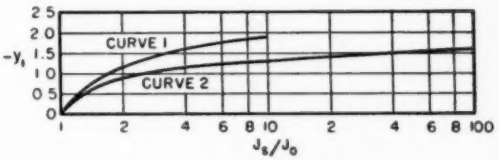


FIG. 5. Variation of position of virtual cathode with anode current. Curve 1 is computed from equations in this paper and curve 2 from Langmuir's curves. y_1 is the normalized coordinate of the cathode, J_0 is the anode direct-current density, and J_s is the saturation-current density.

† This is not an approximate relation and may be shown to be always true by replacing v_0 by $Y=v_0^3/\mu^3$ in Eq. (13).

quency to the plasma frequency at the virtual cathode. Thus ϵ plays an important role and determines whether the process is to be classified as high or low frequency. For electron guns as used in microwave tubes, the process is definitely one of high frequency, as will be shown later. The initial conditions that J must satisfy will now be given.

1.5 Initial Conditions for Current Density J

As mentioned previously, the initial conditions are different for the cases of ac impedance (Fig. 2) and fluctuations (Fig. 3). They will be treated separately.

1.5.1 Case 1: ac Impedance

At $y=y_1$ the initial conditions are

$$J=J_1, \tag{34a}$$

$$J_y=0. \tag{34b}$$

The second initial condition corresponds to the statement that the ac density is zero at the cathode. This will be so, since the density at the cathode, which is the source of electrons, is determined solely by conditions inside the metal and the image forces. These correspond to forces that are much greater than those applied by external means.

From Eq. (25)

$$\begin{aligned} &a_0+a_1y_1+a_2y_1^2=J_1, \\ &a_1+2a_2y_1=0. \end{aligned} \tag{35}$$

Equations (35), (26), and (27) may now be solved for the coefficients and I. The ac voltage drop, V, may be computed as follows:

$$V=-\int_0^d Edx=\frac{J_0D}{j\omega\epsilon_0}\int_{y_1}^{y_2}(J-I)dy. \tag{36}$$

The ac impedance

$$Z_{ac}=-(V/I). \tag{37}$$

While the problem of the calculation of ac impedance has been solved in principle, it is not possible to obtain results for a particular case without lengthy numerical computations. Only the case of $y_1=0$ will be considered here.

It is of interest to digress for a moment to observe how Eq. (27) may be used to solve the complete-space-charge diode of Llewellyn.[11] Here, the dc relation Eq. (17) holds up to $y=0$ at the cathode. The values of c_1, c_2, and I are determined solely by the initial condition $J=J_1$ at $y=0$ and J_1 is finite. Thus, for J to be finite at $Z=0$, it follows immediately that

$$\begin{aligned} c_1&=-\frac{2I}{jB_{12}}, \\ c_2&=-\frac{2I}{j^2B_{12}^2}. \end{aligned} \tag{38}$$

[11] See reference 4, p. 45. Here the thermal-velocity spread is entirely neglected, giving for this case $Z_m=0$.

Hence,

$$J=\frac{2I}{j^2B_{12}^2}\left[\frac{1-jB_{12}Z\exp(-jB_{12}Z)-\exp(-jB_{12}Z)}{Z^2}\right]. \tag{39}$$

Using the initial condition in Eq. (39),

$$I=J_1. \tag{40}$$

It will also follow automatically that at $Z=0$, v and E are both zero. It should be noted that Llewellyn was forced to assume that $v=0$ or $E=0$ at the cathode, but actually this follows automatically from Maxwell's equations for a complete-space-charge region.

1.5.2 Case 1A: Calculation of ac Impedance for $y_1=0$

Here

$$\begin{aligned} a_0&=J_1, \\ a_1&=0, \\ I&=[1-(9/2)\epsilon^2+j3\epsilon]J_1, \\ a_2&=\frac{-j\theta}{12\mu^{5/2}}J_1. \end{aligned} \tag{41}$$

It is to be noted that I is not equal to J_1 as in Llewellyn's case.

The value of the other constants are

$$c_1=-\frac{2J_1}{j3\theta}[1-(9/2)\epsilon^2+j3\epsilon](1+p)\exp j4.13\epsilon,$$

$$c_2=\frac{2J_1}{9\theta^2}[1-(9/2)\epsilon^2+j3\epsilon](1+q)\exp j4.13\epsilon, \tag{42}$$

$$p=-j\epsilon\frac{(4.13+4.25\epsilon^2)}{1-(9/2)\epsilon^2+j3\epsilon}, \tag{43}$$

$$q=\frac{\epsilon}{1+j3\epsilon-(9/2)\epsilon^2}[4.35\epsilon-17.2\epsilon^3+j(15\epsilon^2-4.13)]. \tag{44}$$

The modulation at the anode and ac impedance is given by

$$\begin{aligned} J_2&=-\frac{2}{9}\mu J_1\frac{1-(9/2)\epsilon^2+j3\epsilon}{\epsilon^2} \\ &\times\{1-[jB_{m2}(1+p)+1+q]\exp j(4.13\epsilon-B_{m2})\}, \\ v_2&=-\frac{2}{9}\mu J_1\frac{1-(9/2)\epsilon^2+j3\epsilon}{\epsilon^2} \\ &\times\left\{1-\frac{2}{jB_{m2}}+\frac{1}{jB_{m2}}[jB_{m2}(1+p)+2+2q]\right. \\ &\left.\times\exp j(4.13\epsilon-B_{m2})\right\}, \end{aligned} \tag{45}$$

where $B_{m2}=3\theta(Z_2-Z_m)$, the transit time in the Child's law region.

$$Z_{ac}=Z_{ac}^{(1)}+Z_{ac}^{(2)},$$

$$Z_{ac}^{(1)}=\frac{d}{j\omega\epsilon_0}\left\{1+\frac{6}{j^3B_{m2}^3}\{-(jB_{m2}+2)\right.$$

$$\left.\times\exp[j(4.13\epsilon-B_{m2})]+2-jB_{m2}\}\right\}, \quad (46)$$

$$Z_{ac}^{(2)}=-\left(\frac{kT_0}{m}\right)^{\frac{1}{2}}\frac{G\epsilon}{j\epsilon_0\omega^2}+\frac{6d}{j\omega\epsilon_0B_{m2}^3}$$

$$\times\{(-jB_{m2}p-p-q)\exp[j(4.13\epsilon-B_{m2})]$$

$$+p(j4.13\epsilon+1)+4.13\epsilon+q+j1.35B_{m2}\mu^{\frac{1}{2}}\}$$

$$G=\frac{2.5-j0.435\epsilon}{1-(9/2)\epsilon^2+j3\epsilon}=\frac{2.5-12.5\epsilon^2+j\epsilon(-8+2\epsilon^2)}{1+(81/4)\epsilon^4}. \quad (47)$$

For small ϵ, $Z_{ac}^{(1)}$ reduces to Llewellyn's formula. For large B_{m2} these formulas may be simplified somewhat. It may be noted that for large B_{m2} there will always be a resistive component, while the Llewellyn formula predicts a purely capacitative component. Physically, this is to be expected in view of the velocity spread. The value of the impedance at high frequencies, Z_∞, is

$$Z_\infty=\left[-\left(\frac{kT_0}{m}\right)^{\frac{1}{2}}\frac{G\epsilon}{\omega}+d\right]\frac{1}{j\omega\epsilon_0}. \quad (48)$$

It is seen that for $\epsilon^2>0.2$ there is always a positive resistive component of the impedance. Also, it is possible to conceive values of ϵ, ω, T_0, and d for which Z_∞ has a considerable resistive component. This could possibly be checked experimentally by methods similar to those employed by Robertson.[12]

1.5.3 Case 2: Fluctuations in Diodes

At $y=y_1$ the initial conditions are

$$J=J_1, \quad (49)$$

and the circuit equation is

$$V=IZ_L, \quad (50)$$

in which V is given by Eq. (36).

Sufficient conditions are now known to determine the constants in Eqs. (26) and (27). Unfortunately, the work required to compute the fluctuations for all values of Z_L is very great. Thus, only the case of the open circuit, $Z_L=\infty$ and $I=0$, will be discussed here. This is the case of interest for the diode region in the electron gun of the traveling-wave tube. Furthermore, it is seen from Eqs. (27) and (29) that for large transit angles, J_2 and v_2 will not depend on I. Hence, for large transit angles the open-circuit case will be a good approximation for all diode terminations.

For the open-circuit case, $I=0$, the constants are given as follows:

$$\frac{a_0}{J_1}=\frac{1}{1+y_1(a_1/a_0)+y_1^2(a_2/a_0)},$$

$$\frac{a_1}{a_0}=-\frac{1-4.5\epsilon^2+j3\epsilon}{1.41(1+j3\epsilon)},$$

$$\frac{a_2}{a_0}=-\frac{1}{5.65}\left[2.12j\epsilon+(4+j3\epsilon)\frac{a_1}{a_0}\right]. \quad (51)$$

From these the current modulation J_2 and velocity modulation v_2 at the anode can be computed as follows:

$$\frac{J_2}{J_1}=\left(\frac{1}{Z_2}\frac{c_1}{J_1}+\frac{1}{Z_2^2}\frac{c_2}{J_1}\right)\exp-jB_{m2}, \quad (52)$$

$$\frac{v_2}{J_1}=-\left(\frac{Z_2}{jB_{m2}}\frac{c_1}{J_1}+\frac{2}{jB_{m2}}\frac{c_2}{J_1}\right)\exp-jB_{m2}, \quad (53)$$

and

$$\frac{c_1}{J_1}=\left(2.12\frac{J_m}{J_1}+j3.9\epsilon\frac{v_m}{J_1}\right)\exp j4.13\epsilon,$$

$$\frac{c_2}{J_1}=\left(-1.12\frac{J_m}{J_1}-j4.13\epsilon\frac{v_m}{J_1}\right)\exp j4.13\epsilon,$$

$$\frac{J_m}{J_1}=\left(1+1.13\frac{a_1}{a_0}+1.28\frac{a_2}{a_0}\right)\frac{a_0}{J_1},$$

$$\frac{v_m}{J_1}=1.11\left(\frac{J_m}{J_1}+\frac{1}{j3.9\epsilon}\frac{J_{ms}}{J_1}\right),$$

$$\frac{J_{ms}}{J_1}=3.35\left(\frac{a_1}{a_0}+2.26\frac{a_2}{a_0}\right)\frac{a_0}{J_1}. \quad (54)$$

Also at the virtual cathode

$$J_c=a_0$$

$$v_c=\frac{0.61+2.75\epsilon^2}{9\epsilon^2-j3\epsilon}\frac{a_0}{J_1}. \quad (55)$$

To convert to unnormalized units (denoted by a subscript u) one uses the following:

$$J_u=(J/J_1)J_{1u},$$

$$v_u=\left(\frac{v}{J_1}\right)\frac{J_{1u}}{J_0}\left(\frac{kT_0}{m}\right)^{\frac{1}{2}}1.65, \quad (56)$$

and

$$(J_{1u}S)^2=2eJ_0Sdf, \quad (57)$$

where S is the cross-sectional area of the beam under consideration.

[12] S. D. Robertson, Bell System Tech. J. 28, 619 (1949).

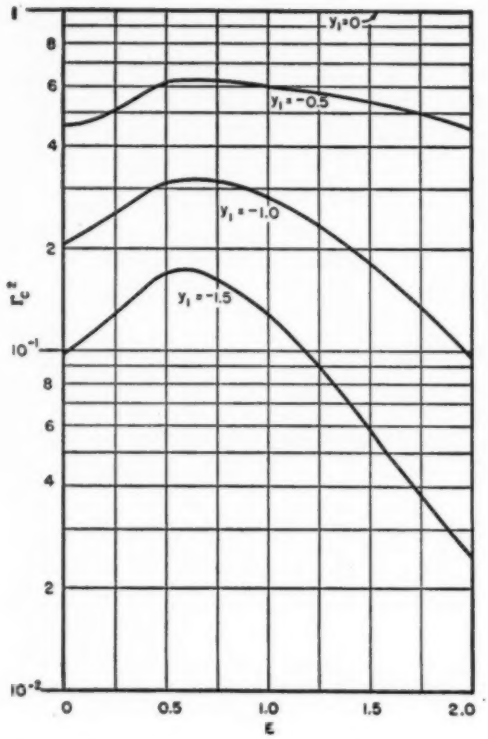


FIG. 6. Current fluctuation at the virtual cathode as a function of frequency, $\epsilon=0.47\ \omega/\omega_p$. $\langle J_c S^2\rangle_{\Delta f}=\Gamma_c^2 2eI_0 df$.

Thus, with the use of Eqs. (52) through (57), one may compute the value of all fluctuating-field quantities. The corresponding equations derived from the RPHP theory are: At the virtual cathode

$$v_u=0.655(kT_0/m)^{\frac{1}{2}}J_{1u}/J_0, \tag{55a}$$

and at the anode[6]

$$v_{2u}=-0.655\left(\frac{kT_0}{m}\right)^{\frac{1}{2}}\frac{J_{1u}}{J_0}\exp -j3.9\epsilon Z_2, \tag{53a}$$

$$J_{2u}=j\frac{2.36\epsilon}{Z_2}(0.655)J_{1u}\exp -3.9\epsilon Z_2. \tag{52a}$$

To convert from Pierce's notation to the one used here, one uses the fact that $B_{\beta 2}$ (transit angle from virtual cathode to anode) $=3.9\epsilon Z_2$. An important deduction from the RPHP theory is that J_{2u} and v_{2u} are 90 degrees out of phase and that

$$J_{2u}/v_{2u}=-j2.36\epsilon J_0/Z_2(kT_0/m)^{\frac{1}{2}}. \tag{58}$$

This ratio is still true for large Z_2 in our theory, as can be seen from Eqs. (52) and (53). Thus for large Z_2 it is necessary only to compute v_2.

Let the following noise factors be defined: At the virtual cathode let

$$\langle (J_{cu}S)^2\rangle_{\Delta f}=\Gamma_c^2 2eI_0 df,$$

$$\langle v_{cu}{}^2\rangle_{\Delta f}=q_c^2\frac{e}{m}\frac{kT_0}{I_0}df,$$

I_0 be the total beam current, and Γ_2^2 and q_2^2 be the corresponding factors at the anode. These noise factors have the following values in the RPHP theory:

$$q_c^2=q_2^2=0.86$$
$$\Gamma_2^2=2.2(\epsilon/Z_2)^2.$$

These noise factors have been plotted *versus* ϵ in Figs. 6 through 9. It is seen that the noise factors decrease as y becomes more negative and that they vary with y, ϵ, and Z_2. Furthermore, there is no question as to the coherence of the various fluctuations, since they are all derived from the shot-noise fluctuations at the cathode. The phase angle between the current and velocity fluctuations at the anode (θ_2) and virtual cathode (θ_c) is plotted against ϵ in Fig. 10. It is seen that θ_2 is almost 90 degrees for large Z_2. However, it is not exactly 90 degrees and does deviate about 3 degrees.

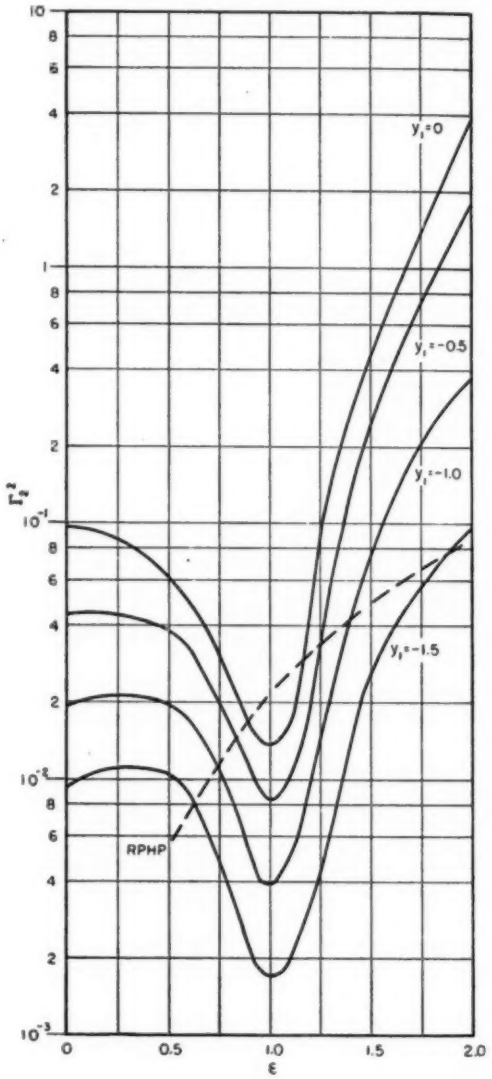


FIG. 7. Current fluctuation at anode plotted against frequency $\epsilon=0.47\ \omega/\omega_p$, $\langle J_2 S^2\rangle_{\Delta f}=\Gamma_2^2 2eI_0 df$, and $Z_2=10$. The dashed line is for the RPHP theory.

1.6 Conclusions

It is possible to work out a theory of ac impedance and fluctuations in diodes by accounting for the thermal-velocity spread by the hydrostatic pressure method of Hahn.

The equations are difficult to solve, and it becomes necessary to make certain approximations. Comparisons of these theories and those of Llewellyn for diode impedance, and those of Rack-Peterson (RP) for fluctuations are made. It is shown that the high frequency impedance always has a resistive component, which may be large in some cases, especially in diodes of small spacing and low anode currents. In general, the fluctuations depend on more parameters and may be greater than predicted by the RP theory. For a given dc condition everything is determined by ϵ, which is proportional to the ratio of the applied frequency to the plasma frequency at the virtual cathode. ($\epsilon=0.47\ \omega/\omega_p$).

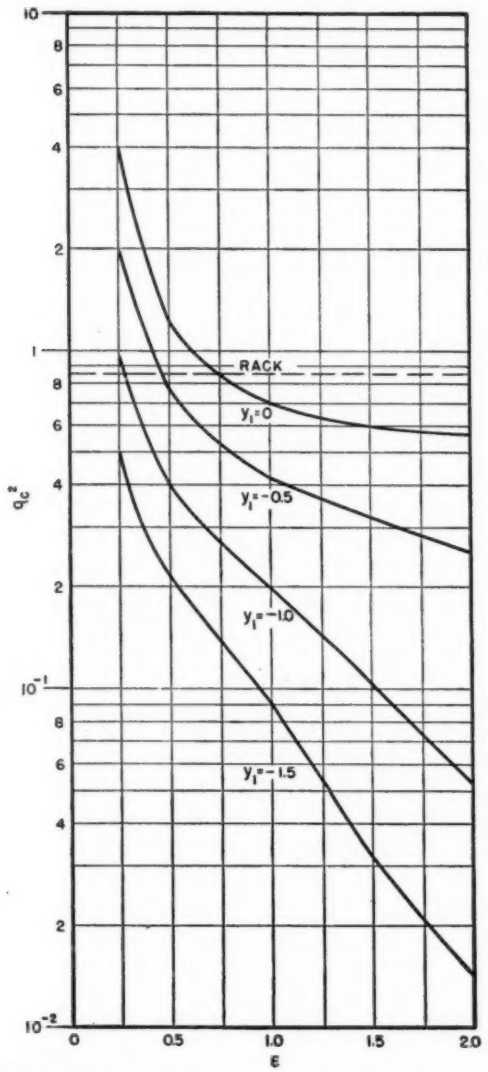


FIG. 8. Velocity fluctuation at virtual cathode *versus* frequency. $\epsilon=0.47\ \omega/\omega_p$ and $\langle v_c^2\rangle_{kn}=q_c^2ekT_0df/mI_0$.

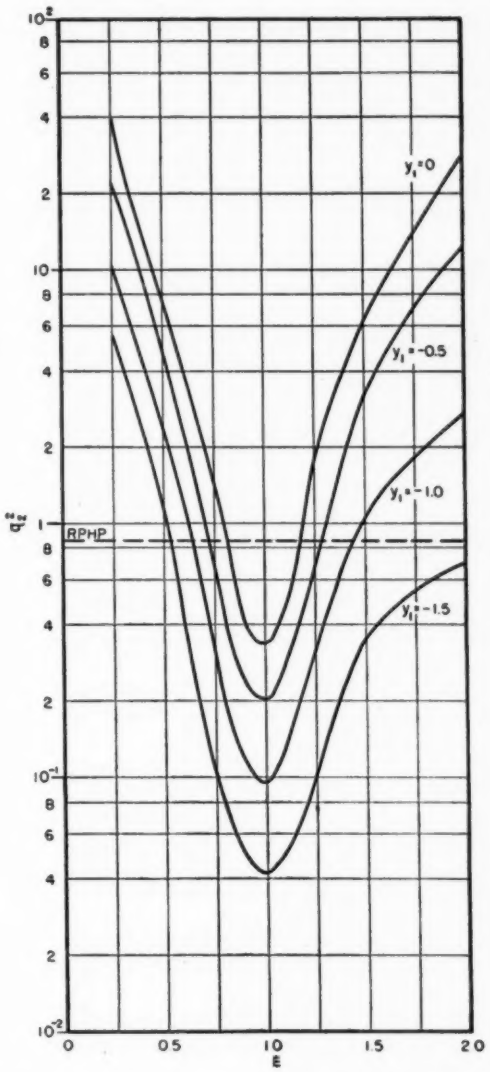


FIG. 9. Velocity fluctuation at anode as a function of frequency. $\epsilon=0.47\ \omega/\omega_p$, $\langle v_2^2\rangle_{kn}=q_2^2ekT_0df/mI_0$, and $Z_2=10$.

2. AC PROPERTIES OF DRIFT TUBE

2.1 General

The propagation of space-charge waves in an electron beam, neglecting the effect of the thermal-velocity spread, has been discussed by Hahn[7] and Ramo.[8] The effect of the thermal-velocity spread will be discussed here.

In this case the beam velocity v_0 is a constant $=v_{0d}$. The subscript, d, will denote the value of quantities in the drift region. In this region let the normalizing parameter $\alpha=v_{0d}$. Hence,

$$D=(2/9)^{\frac{1}{2}}v_{0d}/\omega_d,$$
$$\theta=(2/9)^{\frac{1}{2}}\omega/\omega_d, \tag{59}$$

where ω_d is the plasma frequency in the drift region.

The differential equation for J now reduces to

$$(1-\mu_d^2)J_{yy}+j2\theta J_y+(2/9-\theta^2)J=(2/9)I, \tag{60}$$

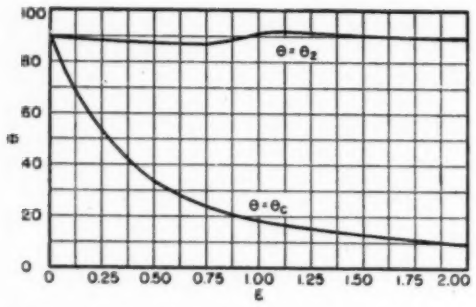


FIG. 10. Phase angle between current and velocity fluctuations at the virtual cathode and anode plotted against frequency. $\epsilon=0.47\,\omega/\omega_p$, θ_2=angle at anode, θ_c=angle at virtual cathode, and $Z_2=10$.

of which the solution is

$$J=\frac{I}{1-(\omega/\omega_d)^2}+a_1\exp(-\Gamma_1 x)+a_2\exp(-\Gamma_2 x), \quad (61)$$

$$\Gamma_1=\frac{j\gamma_0}{1-\mu_d^2}\left\{1+\left[\left(\frac{\omega_d}{\omega}\right)^2(1-\mu_d^2)+\mu_d^2\right]^{\frac{1}{2}}\right\}, \quad (62)$$

$$\Gamma_2=\frac{j\gamma_0}{1-\mu_d^2}\left\{1-\left[\left(\frac{\omega_d}{\omega}\right)^2(1-\mu_d^2)+\mu_d^2\right]^{\frac{1}{2}}\right\}, \quad (63)$$

$$\gamma_0=\omega/v_{0d}, \quad (64)$$

$$\mu_d^2=kT_0/mv_{0d}^2. \quad (65)$$

2.2 Modes in an Infinite Beam

a_1 and a_2 are arbitrary constants to be determined by the initial conditions. There are two possible modes in an infinite beam.

2.2.1 a_1 and $a_2=0$, $I\neq 0$; No Propagation of Space-Charge Waves

This is a mode where there is no propagation. The density modulation is zero, while the ac field and convection-current density are constant along the x direction.

2.2.2 $I=0$; a_1 and $a_2\neq 0$; Propagation of Space-Charge Waves

This is the mode where there is propagation. The case of interest is that for which I vanishes and $\mu_d\ll 1$. Hence,

$$\Gamma_1=j\gamma_0(1+\delta_d), \quad (66)$$

$$\Gamma_2=j\gamma_0(1-\delta_d), \quad (67)$$

and the correction for the space-charge-wave number is

$$\delta_d=\frac{\omega_d}{\omega}\left(1+\frac{kT_0}{mv_{0d}^2}\frac{\omega^2}{\omega_d^2}\right)^{\frac{1}{2}}. \quad (68)$$

Solving for a_1 and a_2 in terms of the current and velocity fluctuations at the beginning of the drift region, the current modulation $J_{3\mu}$ and velocity modulation $v_{3\mu}$ along the drift region is

$$J_{3\mu}=-\exp[-j\gamma_0(x-x_2)]\times jv_{2\mu}\frac{\rho_{0d}}{\delta_d}\sec\phi_0\sin(\phi_0-\phi), \quad (69)$$

$$v_{3\mu}=\exp[-j\gamma_0(x-x_2)]v_{2\mu}\sec\phi_0\cos(\phi_0-\phi), \quad (70)$$

$$\rho_{0d}=\text{dc density in drift region}, \quad (71)$$

$$\tan\phi_0=-\frac{J_{2u}\delta_d}{jv_{2u}\rho_{0d}}=1.4\frac{\omega'_d}{\omega_d}\left(1+\frac{kT_0}{mv_{0d}^2}\frac{\omega^2}{\omega_d'^2}\right)^{\frac{1}{2}}, \quad (72)$$

which follows from Eqs. (58) and (74).

$$\phi=\gamma_0\delta_d(x-x_2). \quad (73)$$

These expressions are valid for an infinite electron beam. For the practical case of a circular electron beam of radius b in a circular drift tube of radius a, one may account for the finite geometry by applying the correction used in the case of no thermal-velocity spread. This[8] is accomplished by reducing the plasma frequency ω_d by a factor given by

$$\omega_d'=\omega_d/[1+(\tau/\gamma_0)^2]^{\frac{1}{2}}. \quad (74)$$

τ/γ_0 is a function of $\gamma_0 b$ and a/b only for small densities. A plot of τ/γ_0 *versus* $\gamma_0 b$ is given in Fig. 11.

According to Eqs. (69) and (70), J_{3u} and v_{3u} are periodic functions of distance of period $2\pi/\gamma_0\delta_d$. The periodicity of J_{3u} has been observed experimentally by several investigators. The effect of temperature is to decrease this period.

3. TRAVELING WAVE-TUBE AMPLIFIER REGION

3.1 General

The usual theory of the traveling wave-tube amplifier, such as that of Pierce,[6] depends on a knowledge of the circuit and ballistic equations. According to Pierce and using his notation, the circuit equation for a filamentary electron beam is in unnormalized units

$$V=\left[\frac{\Gamma\Gamma_1(E^2/B^2P)}{2(\Gamma^2-\Gamma_1^2)}-\frac{j\Gamma}{\omega c_1}\right]i, \quad (75)$$

where $i=JS$, $E=\Gamma V$, and V is voltage.

The ballistic equation, which includes the effect of thermal-velocity spread, is[9]

$$i=\frac{I_0}{2V_0}\frac{jB_e\Gamma V}{(jB_e-\Gamma)^2-\mu_d^2\Gamma^2}, \quad (76)$$

where I_0 is the total beam current, which equals $-J_0S$, and B_e is the electron-beam wave number.

$$v=-\frac{v_{0d}}{2V_0}\frac{(jB_e-\Gamma)\Gamma V}{(jB_e-\Gamma)^2-\mu_d^2\Gamma^2}.$$

Comparing Eqs. (75) and (76), one has the following equation for Γ:

$$(jB_e-\Gamma)^2=\frac{j2B_e\Gamma^2\Gamma_1C^3}{(\Gamma_1^2-\Gamma^2)}+\Gamma^2\left(4QC+\frac{\mu_i^2}{C^2}\right), \quad (77)$$

where C is the Pierce coupling parameter and QC is the Pierce space-charge parameter.

$$\begin{aligned} C^3&=E^2I_0/B^2P8V_0, \\ Q&=B_e/\omega c_1(E^2/B^2P). \end{aligned} \quad (78)$$

Only the case of synchronism and no loss will be considered here. Following Pierce,[6] page 113, let

$$\Gamma=jB_e-j\delta B_eC; \quad \delta\ll 1$$

$\Gamma_1=jB_e$ for synchronism. Hence,

$$\delta^3-[4QC+(\mu_d^2/C^2)]\delta-1=0. \quad (79)$$

For small space charge, QC may be approximately computed by associating the second term of Eq. (75) as because of the propagation of space-charge waves in the electron beam independently of the external circuit. Thus,

$$-\frac{\omega C_1}{j\Gamma}=\frac{I_0}{2V_0}\frac{jB_e\Gamma}{(jB_e-\Gamma)^2-\mu_d\Gamma^2}, \quad (80)$$

and the value of Γ is given by Eq. (66). Thus, approximately

$$\omega C_1/B_e=I_0/2V_0(\omega/\omega_h)^2, \quad (81)$$

where ω_h is the plasma frequency in this region. Hence,

$$4QC=(\omega_h/\omega)^2 1/C^2. \quad (82)$$

A more rigorous solution would be obtained by utilizing the expression of Eq. (76) in the various field theories of the traveling wave tube. This will not be carried out here.

It is of interest to digress for a moment to observe that the space-charge parameter of Pierce for small QC may be obtained by modifying the ballistic equation to include a polarization term. Neglecting thermal-velocity spread, the equation of motion is

$$m dv/dt=e[E-(P/\epsilon_0)]; \quad (83)$$

P, which is the polarization, equals $n_0e\zeta$. ζ is the ac displacement of an electron. The reason for this is as follows: The acceleration of an electron is proportional to the force, F, acting on the electron under consideration. E, however, includes the force resulting from the electron under consideration. Thus, in order to obtain F, one must subtract the force resulting from the electron, which is approximated on the average by the electric field in a small slice bounded by infinite parallel planes around the electric field.

This force equals P/ϵ_0. Hence $F=E-(P/\epsilon_0)$. The alternating current

$$i=\frac{\partial P}{\partial t}=\frac{I_0}{2V_0}\frac{jB_e\Gamma V}{(jB_e-\Gamma)^2+(\omega_h/\omega)^2B_e^2}. \quad (84)$$

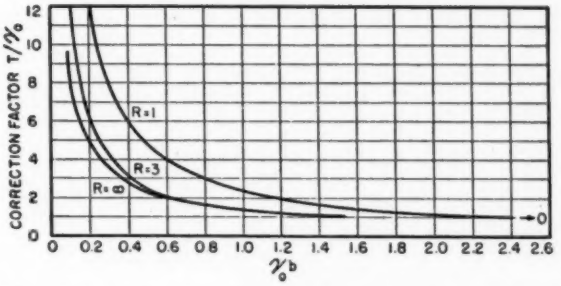


FIG. 11. Correction factor to the plasma frequency in the drift tube for beam and drift-tube geometries. R is the ratio of the radius of the drift tube to the radius of the beam. $\gamma_0=\omega/v_{0d}=2\pi f/$ beam velocity. $\omega_d'/\omega_d=1/[1+(\tau/\gamma_0)^2]^{\frac{1}{2}}$, where ω_d is the plasma frequency in the drift tube. and ω_d' is the plasma frequency corrected for beam and drift-tube geometries.

Combining this with the circuit equation (75) and with $C_1=\infty$, one has

$$(jB_e-\Gamma)^2=\frac{j2B_e\Gamma^2\Gamma_1C^3}{(\Gamma_1^2-\Gamma^2)}-\left(\frac{\omega_h}{\omega}\right)^2B_e^2. \quad (85)$$

Now since Γ can be approximated by jB_e for small space charge at synchronism and using Eq. (82), it is seen that Eq. (85) is equivalent to Eq. (76) for $\mu_d=0$. Thus, in this case the Pierce space-charge correction may be accounted for on purely ballistic grounds.

3.2 Computation of Gain and Noise Figure

Let

$$\Omega_s=4QC=\left(\frac{\omega_h}{\omega}\right)^2\frac{1}{C^2},$$

$$\Omega_t=\frac{\mu_d^2}{C^2}=\frac{kT_0}{mv_{0d}^2}\frac{1}{C^2},$$

$$\Omega=\Omega_s+\Omega_t, \quad (86)$$

then

$$\delta^3-\Omega\delta-1=0. \quad (79a)$$

The three solutions of this cubic correspond to three propagating modes of voltage amplitudes V_m, $m=1, 2, 3$. The distribution among the three modes is determined by the initial conditions, which are

$$\sum E_m=E_i-\frac{j}{\omega C_1}\sum\Gamma_m^2i_m, \quad (87)$$

$$\sum\frac{jB_e-\Gamma_m}{(jB_e-\Gamma_m)^2-\mu_d\Gamma_m^2}E_m=-\frac{2V_0}{v_0}v_{3w}, \quad (88)$$

$$\sum\frac{E_m}{(jB_e-\Gamma_m)^2-\mu_d\Gamma_m^2}=\frac{2V_0}{jB_eI_0}J_{3w}S. \quad (89)$$

Only the case of $\mu_d\ll 1$ will be treated here. With

this condition these equations reduce to

$$\sum_1^3 e_m\left(1-\frac{2\Omega_t C}{\delta_m}-\frac{\Omega}{\delta_m^2}\right)=E_i, \tag{90}$$

$$\sum\frac{e_m}{\delta_m}=-j\frac{2V_0B_eC}{v_{0d}}v_{3u}, \tag{91}$$

$$\sum\frac{e_m}{\delta_m^2}=-\frac{2V_0}{jB_eI_0}B_e^2C^2J_{3u}S, \tag{92}$$

$$e_m=E_m/1-\frac{2\Omega_t C}{\delta_m}-\frac{\Omega_t}{\delta_m^2}. \tag{93}$$

Thus, the gaining wave E_1 (δ_1 is complex) is given by

$$E_1=\left(1-\frac{2\Omega_t C}{\delta_1}-\frac{\Omega_t}{\delta_1^2}\right)\frac{E_i+j\dfrac{2V_0B_eC}{v_{0d}}(-\delta_1-2\Omega_tC)v_{3u}-\dfrac{2V_0B_e^2C^2}{jB_eI_0}\left(\dfrac{1}{\delta_1}+\Omega\right)J_{3u}S}{\left(1-\dfrac{\delta_2}{\delta_1}\right)\left(1-\dfrac{\delta_3}{\delta_1}\right)}. \tag{94}$$

The noise figure F is given by

$$F-1=\frac{\left|j\dfrac{2V_0B_eC}{v_{0d}}(-\delta_1-2\Omega_tC)v_{3u}-\dfrac{2V_0B_e^2C^2}{jB_eI_0}\left(\dfrac{1}{\delta_1}+\Omega\right)J_{3u}S\right|}{|E_i|^2}. \tag{95}$$

Now from Eq. (70) and Fig. 9

$$|v_{3u}|^2=q_2^2\frac{e}{m}\frac{kT_0}{J_0S}df\sec^2\phi_0\cos^2(\phi_0-\phi),$$

$$J_{3u}=-j\frac{\rho_{0d}}{\delta_d}\tan(\phi_0-\phi)v_{3u}. \tag{96}$$

Also in rms values

$$|E_i|^2=B_e^2\frac{4V_0}{I_0}C^3kT_adf, \tag{97}$$

where T_a is the ambient temperature. Hence,

$$F-1=\frac{q_2^2}{2C}\frac{T_0}{T_a}\frac{\sec^2\phi_0}{1+(\delta_d/C)^2Y^2}\times\left|j(-\delta_1-2\Omega_tC)-Y\left(\frac{1}{\delta_1}+\Omega\right)\right|^2, \tag{98}$$

$$Y=C/\delta_d\tan(\phi_0-\phi),$$

in which T_0 is the cathode temperature. This may be further reduced by noting that

$$\tan\phi_0=1.4\frac{\omega_d'}{\omega_d}\left(1+\frac{kT_0}{mv_{0d}^2}\frac{\omega^2}{\omega_d'^2}\right)^{\frac{1}{2}},$$

$$(\delta_d/C)^2=\Omega, \tag{72}$$

according to Eq. (68). Hence,

$$F-1=\frac{q_2^2}{2C}\frac{T_0}{T_a}\frac{\sec^2\phi_0}{1+\Omega Y^2}\times\left|j(-\delta_1-2\Omega_tC)-Y\left(\frac{1}{\delta_1}+\Omega\right)\right|^2, \tag{99}$$

$$Y=-\frac{1}{\Omega^{\frac{1}{2}}}\tan(B_eC\Omega^{\frac{1}{2}}(x_3-x_2)-\phi_0), \tag{100}$$

$$\sec^2\phi_0=1+1.96\left(\frac{\omega_d'}{\omega_d}\right)^2\left(1+\frac{kT_0}{mv_{0d}^2}\frac{\omega^2}{\omega_d'^2}\right).$$

In Fig. 12, δ_1 has been plotted against Ω. Thus one has all the necessary formulas to compute F.

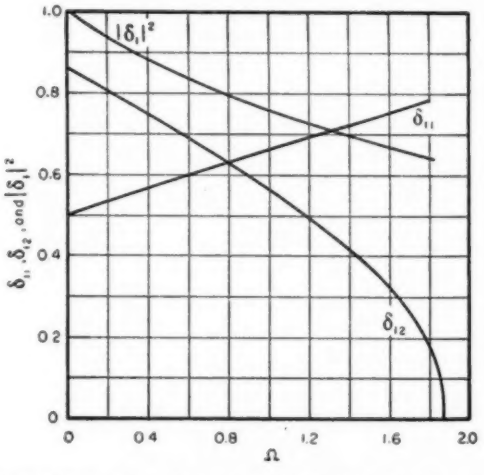


FIG. 12. Relative phase shift δ_{11}, logarithmic gain δ_{12}, and $|\delta_1|^2$, $\delta_1=-\delta_{11}-j\delta_{12}$ are plotted against $\Omega=4QC+kT_0/mv_{0d}^2\times 1/C^2$.

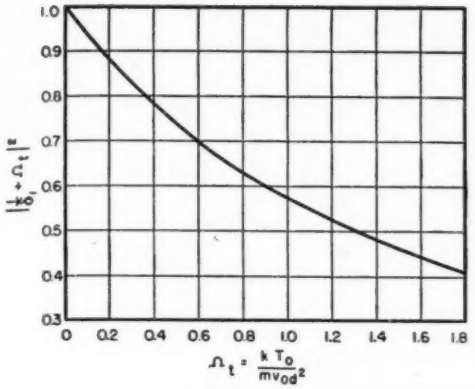


FIG. 13. Reduction in noise figure by thermal-velocity spread in a current-modulated electron beam.

It is seen that the noise figure is a periodic function of the length of the drift tube of period $\pi/B_e C\Omega^{\frac{1}{2}}$. Thus it is possible to vary Y from 0 to ∞ by varying the length of the drift tube. These correspond to the cases where the beam emerges from the drift tube with only velocity modulation or only current modulation, respectively. In general, $\Omega_t C$ is quite small. Hence, Eq. (98) may be rewritten as

$$F-1=\frac{q_2^2}{2C}\frac{T_0}{T_a}\sec^2\phi_0 f_1, \tag{101}$$

where

$$f_1=\frac{\left|j\delta_1+Y\left(\frac{1}{\delta_1}+\Omega\right)\right|^2}{1+\Omega Y^2}. \tag{102}$$

Thus, f_1 measures the effect of the drift tube and the traveling-wave tube while the other factor measures the effect of the diode. In those cases where Ω_s is much larger than Ω_t, the effect of the thermal spread will appear only in the factor q^2. This is probably the case in practice since $\mu_d^2\sim 1\times 10^{-4}$ for a 500-volt beam, and $(\omega_d'/\omega)^2$ will be greater than 10^{-4}. In the case where Ω_s is much less than Ω_t, there will be a reduction in the noise factor because of the thermal-velocity spread, which can be estimated by considering the cases $Y=0$ and $Y=\infty$. For $Y=0$,

$$f_1=|\delta_1|^2, \tag{103}$$

which is plotted in Fig. 12. For $Y=\infty$,

$$f_1=\frac{\left|\frac{1}{\delta_1}+\Omega\right|^2}{\Omega^2}. \tag{104}$$

A plot of $|(1/\delta)+\Omega|^2$ against Ω is given in Fig. 13. Thus, it is seen that there is a reduction in noise factor because of the thermal-velocity spread.

It should be remarked that in actual practice C and Q are not too easy to evaluate. Thus a rigorous theory of the effect of thermal-velocity spread can exist only in a field theory.

The case of $\Omega_t=0$ and $q^2=0.86$ has already been discussed in great detail by others, such as the Stanford group. This will not be discussed here.

A numerical example is as follows:

diode spacing	$=0.635$ cm
beam radius	$=0.064$ cm
drift-tube radius	$=0.192$ cm
T_0	$=1000°$K
dc diode voltage	$=500$ volts
direct-current density	$=63\times 10^{-3}$ amp/cm^2
ω	$=3\times 10^{10}$/cycle
$\gamma_0 b$	$=1.5$
T/γ_0	$=1$
ω_d'	$=0.7\times 10^9$
C	$=0.075$‡
ϵ	$=1.34$
Z_2	$=10$
q_2^2	$=0.6$
Γ_2^2	$=0.04$
Ω_s	$=0.08$
Ω_t	$=0.02$
y_1	$=-1$§

In this case, as seen from Eq. (101), the effect of the thermal-velocity spread will be felt only in the factor q^2. Thus, the RPHP theory in practice will be qualitatively correct. One cannot, however, expect a quantitative agreement between theory and experiment as seen from Fig. 9. For less dense electron beams the RPHP theory will not be qualitatively correct, since the effect resulting from Ω_t will predominate.

4. CONCLUSIONS

It is possible to find the effect of the thermal velocity on noise figure by extending a method of Hahn. This effect can be characterized by the calculation of two quantities, q_2^2 and Ω_t. For a given set of dc conditions in the diode region, q_2^2 is a function only of ϵ, which is proportional to the ratio of applied frequency to the plasma frequency at the virtual cathode. It is this quantity ϵ that determines whether the process is to be characterized as that of low or high frequency. Computations are given for q_2^2 as a function of ϵ. These results agree with those computed from the Rack-Peterson theory only under certain conditions.

For those cases where $\Omega_t \ll 4QC$, the effect of the thermal-velocity spread will be determined solely by q_2^2. Thus, the RPHP theory will still give qualitatively correct results. Experimental agreement with the RPHP theory within a few decibels, however, is not to be expected.

The noise figure F at synchronism is given by

$$F-1=\frac{q_2^2}{2C}\frac{T_0}{T_a}\frac{\sec^2\phi_0}{1+\Omega Y^2}\left|j\delta_1+Y\left(\frac{1}{\delta_1}+\Omega\right)\right|^2, \tag{99a}$$

‡ This was computed from Fig. A6.5 in Pierce's book and is probably too high.

§ This assumes a saturation-current density of 120 ma/cm^2.

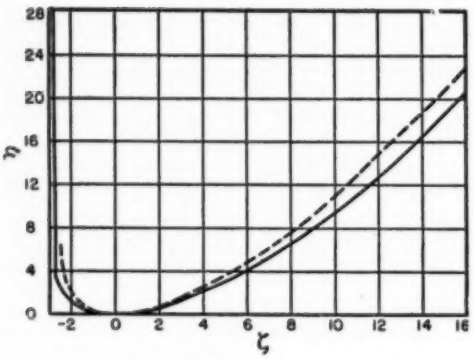


FIG. 14. Potential distribution in diode with thermal velocities according to Langmuir (dashed) and Hahn (solid) curves. $\eta = eV/kT_0$ and $\zeta = 1.58y$.

where q_2^2 is a function of ϵ and is given in Fig. 9,

$$\Omega = 4QC + \frac{kT_0}{mv_{0d}^2}\frac{1}{C^2},$$

$$Y = -\frac{1}{\Omega^{\frac{1}{2}}}\tan[B_e C\Omega^{\frac{1}{2}} l_d - \phi_0], \tag{100}$$

$l_d = (x_3 - x_2)$ = length of drift tube, and δ_1 is given in Fig. 12.

In the course of this work the general theory of the diode, including the effect of the thermal-velocity spread, has also been derived. An experimental check of this theory would be to measure the diode resistance at very high frequencies.

5. ACKNOWLEDGMENT

The aid of Miss Betsy Jones in performing the computations is gratefully acknowledged.

6. APPENDIX

For the exact solution of the dc equation we may start with Eq. (15)

$$v_0^2 = \mu^2\left(1+\log\frac{v_0^2}{\mu^2}\right) + \frac{4}{9}\int_0^y dy \int_0^y \frac{dy}{v_0}. \tag{15}$$

Introducing the new independent variable

$$\rho = \int_0^y \frac{dy}{v_0}, \tag{15a}$$

one obtains

$$v_0^2/\mu^2 = 1 + \log v_0^2/\mu^2 + \frac{4}{9}\int_0^\rho \rho v_0 d\rho \tag{15b}$$

and, differentiating this with ρ, it follows

$$2v_0 v_{0\rho} = 2\frac{\mu^2}{v_0}v_{0\rho} + \frac{4}{9}\rho v_0, \tag{15c}$$

of which the solution is

$$v_0 + \frac{\mu^2}{v_0} - 2\mu = \frac{1}{9}\rho^2, \tag{15d}$$

and

$$v_0 = \mu + \frac{1}{3}\mu^{\frac{1}{2}}\rho\left(1+\frac{1}{36}\frac{\rho^2}{\mu^3}\right)^{\frac{1}{2}} + \frac{1}{18}\rho^2. \tag{15e}$$

The expressions for y and the potential distribution follow from Eq. (15e). Thus,

$$y/\mu^{\frac{1}{2}} = P + 4\left[\left(1+\frac{P^2}{36}\right)^{\frac{1}{2}} - 1\right] + \frac{P^3}{54}, \tag{15f}$$

$$\eta = \frac{eV}{kT_0} = \frac{2}{9}\left[\frac{P^2}{2} + \frac{P^4}{72} + \frac{1}{3}\int_0^P P^2\left(1+\frac{P^2}{36}\right)^{\frac{1}{2}} dP\right], \tag{15g}$$

$$P = \rho/\mu^{\frac{1}{2}}. \tag{15h}$$

ρ and P vary from $-\infty$ to $+\infty$. As ρ approaches $-\infty$, y approaches the asymptotic value $-4\mu^{\frac{1}{2}} = -1.9$. A curve of η *versus* $\zeta(1.58y)$ is shown in Fig. 14, which is compared to the corresponding curve of Langmuir.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 4 APRIL, 1952

Multiple Scattering of Radiation by an Arbitrary Planar Configuration of Parallel Cylinders and by Two Parallel Cylinders*

VICTOR TWERSKY
Mathematics Research Group, Washington Square College, New York University, New York, New York
(Received November 19, 1951)

The formal solution for the scattering of a plane wave by an arbitrary configuration of parallel cylinders previously presented is specialized to consider the case where all the axes lie in the same plane. The scattered wave is expressed as an infinite sum of orders of scattering, the first order being the usual single scattering approximation. It is shown that the far field form of the multiply scattered orders is symmetrical with respect to the plane of the configuration for minimum spacing large compared to wavelength. Hence, for the analogous configuration of bosses on a perfectly conducting plane, departures from the predictions of single scattering theory should occur primarily for the component polarized perpendicular to the elements. The problem of two cylinders is considered in detail, the multiply scattered orders being summed explicitly subject to the above conditions. Approximate solutions for radii very small compared to wavelength are derived, the wavelengths for which the effects of multiple scattering are greatest are investigated, and a criterion for the use of the single scattering hypothesis presented. Solutions for two bosses and a single cylinder parallel to a nonabsorbing plane are also stated. Both the electromagnetic and acoustic cases are treated.

1. INTRODUCTION

In a previous paper,[1] a formal solution for the scattering of a plane wave, $\psi_i=\exp[iKr\cos(\theta+\alpha)]$, by an arbitrary configuration of $N+1$ parallel cylinders of radii a_s was obtained; the usual prescribed boundary conditions were satisfied simultaneously at all surfaces by explicitly taking into account all contributions to the excitation of each cylinder by the radiation scattered by the remaining cylinders. The scattered wave was expressed as an infinite sum of "orders of scattering," or

$$\psi=\sum_{m=1}^{\infty}\psi^m=\sum_{s=0}^{N}\sum_{m}{}^s\psi^m=\sum_{s}\sum_{n=-\infty}^{\infty}{}^sA_nH_n(Kr_s)\times e^{in(\theta_s+\alpha+\pi/2)+i\mathbf{K}\cdot\mathbf{r}_{os}}\sum_{m}{}^sB_n{}^m, \quad (1)$$

where $\Psi=\psi+\psi_i$ is the velocity potential for acoustics or the scalars of the Hertz vector potentials for electromagnetics;[2] the ${}^sA_n$ are the scattering coefficients of the single cylinder (e.g., ${}^sA_n=-J_n(Ka_s)/H_n(Ka_s)$ for $\Psi|_{r_s=a_s}=0$, etc.); H is the Hankel function of the first kind to correspond to outgoing waves for the suppressed time factor $\exp(-i\omega t)$; r_s and θ_s are the coordinates of a field point P in the coordinate frame of the s-th cylinder, the polar axis of all frames being parallel to the x-axis; and $\mathbf{r}_{os}$ is the position of the s-th cylinder from the reference origin.

The ${}^sB_n{}^m$ equal unity for $m=1$, ψ^1 being the first order of scattering (resulting from the excitation of each cylinder by only the incident plane wave), and $\psi_i+\psi^1$ is the usual single scattering approximation. For the higher orders the ${}^sB_n{}^m$, comprising sums of products of the scattering coefficients of the single cylinder and Hankel and trigonometric functions of the geometry, can be written for brevity as the formal product,

$${}^sB_n{}^m=\prod_{\mu=(1)}^{(m-1)}\{\sum_{n^\mu}\sum_{s^\mu}{}'{}^{s^\mu}A_{n^\mu}H_{n^\mu-1_{-n^\mu}}(Kr_{s^\mu-1_{s^\mu}})\times\exp[-i(n^{\mu-1}-n^\mu)(\theta_{s^\mu-1_{s^\mu}}+\alpha+\pi/2)+i\mathbf{K}\cdot\mathbf{r}_{\mu-1_{s^\mu}}]\}, \quad (2)$$

where μ indicates the number of primes distinguishing the indices; $\mathbf{r}_{ss'}$, $\theta_{ss'}$, etc., are the coordinates that locate cylinder s' with respect to cylinder s; and $\sum_{s^\mu}{}'$ indicates that $s^\mu=s^{\mu-1}$ is excluded from the sum.

In the present and in subsequent papers we will specialize this result for application to specific configurations or distributions of general interest. Section 2, in which the general expression is specialized for an arbitrary planar configuration and an "asymptotic" representation obtained, is in its present context merely a step in the derivation of the solution for two cylinders, the simplest configuration and consequently the first to be treated in any detail. However, because of an interesting physical consequence of a symmetry property of the far field form of the multiply scattered waves, which obtains for all planar configurations with minimum spacing large compared to wavelength, this material has been separated and the discussion carried over to the analogous problem of semicylindrical bosses on an infinite perfectly conducting plane. The bosses on the plane provide a simple model to treat scattered reflection from a striated surface, and it is shown that only the component polarized perpendicular to the elements will exhibit departures from single scattering theory to this order of approximation.

* This work was supported in part by Contract AF-19(122)42 with the U. S. Air Force through sponsorship of the Geophysics Research Division of the Air Force Cambridge Research Center, Air Materiel Command. The mathematical details of this paper and further discussion will be found in "Multiple Scattering of Radiation, Part I," Research Report EM-34, Math. Research Group, W.S.C., N.Y.U., 1951, Chapters 2 and 3.

[1] V. Twersky, J. Acoust. Soc. Am. 24, 42 (1952). This paper also contains a representative bibliography of previous work in multiple scattering.

[2] See Appendix.

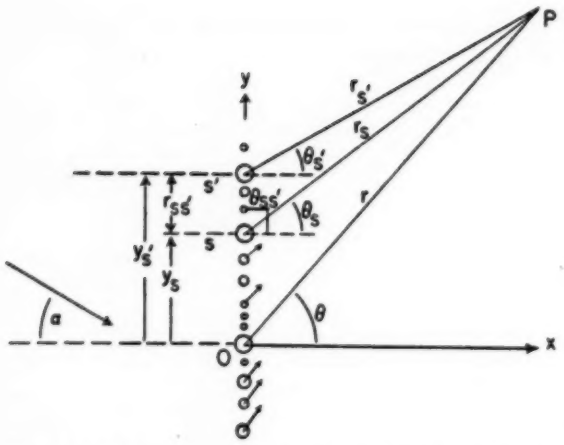


FIG. 1. Arbitrary planar configuration of parallel cylinders.

Section 3 deals with two identical cylinders. For separation large compared to wavelength and small compared to the distance of observation, the far field form of the multiply scattered orders are summed to obtain an explicit representation based on the solution for the single cylinder. Approximate solutions, for radii very small compared to wavelength, are then investigated. Sections 4 and 5 treat two variations of the two-cylinder problem; two bosses on a nonabsorbing plane, and a single cylinder parallel to a nonabsorbing plane.

2. PLANAR CONFIGURATION

Consider a plane wave incident on a configuration of cylinders whose axes lie in the zy-plane parallel to the z-axis as in Fig. 1. The midpoint of the configuration is taken as the origin of the reference coordinates r, θ, z, or x, y, z; the distance from the s-th cylinder to the field point P is $r_s=[x^2+(y-y_s)^2]^{\frac{1}{2}}$, where y_s is the distance of the s-th cylinder from the reference origin. The separation of any two cylinders is $r_{ss'}=y_{ss'}=|y_{s'}-y_s|=\epsilon_{ss'}(y_{s'}-y_s)$, where $\epsilon_{ss'}=\pm 1$ for $s'\gtrless s$, the angle this positive vector makes with the x-axis being $\theta_{ss'}=\epsilon_{ss'}\pi/2$. The phase of the incident wave at the s-th cylinder is $\mathbf{K}\cdot\mathbf{r}_{os}=-Ky_s\sin\alpha$, while the phase difference between the waves originating at cylinders s and s' is $\mathbf{K}\cdot\mathbf{r}_{ss'}=-\epsilon_{ss'}y_{ss'}K\sin\alpha$. Substituting these values into (1) and (2) yields the total scattered wave for a planar configuration

$$\psi=\sum_n i^n\sum_s {}^sA_nH_n(Kr_s)e^{in\theta_s-iKy_s\sin\alpha}$$
$$\times\sum_m\prod_\mu\{\sum_{n_\mu}\exp[i(n^\mu-n^{\mu-1})\pi/2]\sum_{s_\mu}{}' {}^{s_\mu}A_{n_\mu}$$
$$\times H_{n_\mu-1_{-n_\mu}}(Ky_{s_\mu-1_{s_\mu}})\exp[-i\epsilon_{s_\mu-1_{s_\mu}}$$
$$\times[Ky_{s_\mu-1_{s_\mu}}\sin\alpha+(n^{\mu-1}-n^\mu)\pi/2]]\}\exp[in^{(m-1)}\alpha]. \quad (3)$$

For $Kr_s\gg 1$, $Kr_s\gg|n|$, and $Ky_{s_\mu-1_{s_\mu}}\gg 1$, $Ky_{s_\mu-1_{s_\mu}}\gg|n^{\mu-1}-n^\mu|$ (the conditions on the indices to be met by restricting the remaining parameters so that the ${}^sA_n$ approach zero rapidly with increasing n), we can employ $H_\nu(X)\to(2/i\pi X)^{\frac{1}{2}}\exp(iX-i\nu\pi/2)=H_0(X)i^{-\nu}$, where it is understood that H_0 is the asymptotic representation. Substituting into (3), letting $\sin\alpha=A$, and suppressing K for brevity yields the "asymptotic form,"

$$\psi=\sum_n(-1)^n\sum_s {}^sA_nH_0(r_s)e^{in\theta_s-iy_sA}$$
$$\times\sum_m\prod_\mu\{\sum_{n_\mu}\sum_{s_\mu}{}' {}^{s_\mu}A_{n_\mu}H_0(y_{s_\mu-1_{s_\mu}})$$
$$\times\exp[-i\epsilon_{s_\mu-1_{s_\mu}}[y_{s_\mu-1_{s_\mu}}A+(n^{\mu-1}-n^\mu)\pi/2]]\}$$
$$\times\exp[in^{(m-1)}(\alpha+\pi)]. \quad (4)$$

It can be seen that the multiply scattered terms are all symmetrical with respect to the plane of the configuration, i.e., the factor $\exp[in(\theta_s-\epsilon_{ss'}\pi/2]$ they possess assures us that only the sines of an odd multiple of θ or the cosines of an even multiple of θ will appear when the terms of positive and negative n are added. This result can be interpreted as indicating that subject to these conditions the multiple excitations of each cylinder are caused by the waves scattered by the remaining cylinders at angles of $\theta=\pm\pi/2$, or in the plane of the configuration.

We will employ (4) in most future discussions because of the marked simplification it introduces. It should be kept in mind, however, that by its use we consistently neglect terms of the order of, or less than, $-2iH_0/X$ compared to H_0, where $X=(Ky_{ss'})_{\min}$. Hence the multiply scattered orders also possess components symmetrical with respect to the x-axis, but these are small provided that the minimum spacing is large compared to wavelength. Although not necessarily negligible their effects are relatively slight compared to those of the H_0 terms and will be ignored for simplicity although by so doing we introduce discrepancies in regard to order of magnitude of X in the terms retained. The restrictions on X, tacit in the above, are that $X\gg 1$, $X\gg Ka_s$, and these minimize the shielding effects which the present form ignores.

For the analogous configuration of semicylindrical bosses of arbitrary physical constants on an infinite nonabsorbing plane and an incident plane wave $\psi_i(\pi-\theta)$ (the image of the previous $\psi_i\equiv\psi_i(\theta)$ obtained by replacing θ by $\pi-\theta$), the solution is obtained by treating the "image problem," i.e., the previous configuration of cylinders with two incident plane waves which are images in the removed plane. (See treatment of the analogous single scattering problems.[3,4]) Thus for the component polarized perpendicular to the axes of bosses on a perfectly conducting plane (or a rigid plane for acoustics), the total incident wave for the image problem is $\psi_i(\theta)+\psi_i(\pi-\theta)$; the total scattered wave is $\psi''=\psi(\theta_s)+\psi(\pi-\theta_s)$, where $\psi(\theta_s)$ is as in (3) or (4) and $\psi(\pi-\theta_s)$ has θ_s replaced by $\pi-\theta_s$; and the total reflected wave for the boss problem is $\psi_r''=\psi_i(\theta)+\psi''$. For the component polarized parallel to the axes (or

[3] Rayleigh, Phil. Mag. **14**, 350 (1907).
[4] V. Twersky, J. Acoust. Soc. Am. **22**, 539 (1950); **23**, 336 (1951); and J. Appl. Phys. **22**, 825 (1951).

a plane of zero impedance for acoustics), the total incident wave for the image problem is taken as $\psi_i(\theta) - \psi_i(\pi-\theta)$; the total scattered wave is $\psi' = \psi(\theta_s) - \psi(\pi-\theta_s)$; and the total reflected wave for the boss problem is $\psi_r' = \psi_i(\theta) + \psi'$.

Hence as a consequence of the symmetry of the multiply scattered orders for the asymptotic case, we obtain $\psi'^m = 2\psi^m$ while $\psi'^m = 0$, where the ψ^m's are as in (4) with $m > 1$. (For specific configurations we employ (3) to obtain the next order terms in X when $\psi'^m = 0$.) Thus to this order of approximation the perpendicular component will show the effect of multiple scattering while the parallel component is simply the single scattered quantity. It is interesting to note that "anomalies" of single scattering theory which have been observed experimentally with reflection gratings[5] were observed primarily for the perpendicular component (departures being relatively slight or nondetectable for the parallel component), and also that the expression obtained by Rayleigh[3] for reflection from striated surfaces did not predict the behavior of the perpendicular component. (In subsequent papers we will treat the grating and a random planar configuration in detail.)

3. TWO CYLINDERS

The solution of the scattering problem for two identical cylinders as in Fig. 2 is obtained by substituting $y_{s\mu-1,s\mu} = 2b$, $y_s = sb$, and $\epsilon_{s\mu-1,s\mu} = s^\mu = \pm 1$ into (3) for the planar configuration and adding ψ_i. Similarly the asymptotic form for the total scattered wave is obtained by introducing these values into (4). For $r \gg 2b$ we may employ $r_s \approx r - sb\sin\theta$, $\theta_s \approx \theta - (sb/r_s)\cos\theta \to \theta$, so that on neglecting $sb\sin\theta \ll r$ in the denominator of the square root in $H_0(r_s)$, we obtain $H_0(r_s) \to H_0(r) \times \exp(-isKb\sin\theta)$. To this order of approximation, and employing $\gamma = Kb(\sin\theta + \sin\alpha) = T + A$, $H_0(Kr) = H_0$ and $H_0(2Kb) = H$ for brevity, the total scattered wave is

$$\psi = H_0 \sum_n (-1)^n A_n e^{in\theta} \sum_{s=\pm 1} e^{-is\gamma} \sum_m H^{m-1} \prod_{\mu=(1)}^{(m-1)} \times \{\sum_{n\mu} A_{n\mu} \sum_{s\mu}{}' \exp[-is^\mu[2A + (n^{\mu-1} - n^\mu)\pi/2]]\} \times \exp[in^{(m-1)}(\alpha+\pi)], \quad (5)$$

which can be put in a form in which all the even and all the odd multiply scattered orders can be summed explicitly.

Thus on summing the terms for positive and negative values of n, we obtain after a certain amount of bookkeeping that

$$\psi = 2H_0\{C\cos\gamma + H\Gamma_e + H^2C_r\Gamma_0 + H^3C_r{}^2\Gamma_e \cdots\}. \quad (6)$$

For brevity we have employed $C = \sum \epsilon_n A_n \cos n(\theta+\alpha)$, $(n=0$ to $n^* \approx X+2$; $X = Ka$ or $X = \nu Ka$, where ν is the index of refraction; $\epsilon_n = 1$ if $n=0$, $\epsilon_n = 2$ if $n>0)$;

$$\Gamma_e = \tfrac{1}{2}\sum \epsilon_n A_n \cos n(\theta+\pi/2) \sum \epsilon_{n'} A_{n'} \cos n'(\alpha - \pi/2) \times [(-1)^{n+n'} e^{-i\gamma'} + e^{i\gamma'}] = \mathcal{C}_e \cos\gamma' + i\mathcal{C}_0 \sin\gamma',$$

where $\mathcal{C}_e$ contains all terms in which n and n' have even parity, $\mathcal{C}_0$ contains those of odd parity, and where $\gamma' = \gamma - 2A = T - A$; $C_r = \sum \epsilon_n A_n (-1)^n$; and Γ_0 is Γ_e with α replaced by $-\alpha$. The presence of the various terms in the multiply scattered orders may be interpreted as follows: the factor $H^{m-1} = [H_0(2Kb)]^{m-1}$ indicates the number of times the singly scattered wave giving rise to the m'th order of scattering has transversed the spacing $2b$; the "single cylinder factor" C with $\alpha = \pi/2$, in Γ_e and Γ_0, indicates that the excitation reaches the cylinders in the plane of their axes; the factor C with $\theta = -\pi/2$ in Γ_e in the even orders indicates that each cylinder scatters these order in the direction of the remaining cylinder whose first order of scattering initiated these terms; the factor C with $\theta = \pi/2$ in Γ_0 in the odd orders indicates that these are scattered in the same direction as its own first order of scattering; while the intervening factors or multiples of C_r corresponding to the many "internal reflections," indicate that the radiation involved is the back scattered radiation of each cylinder. The quantity $\gamma = Kb(\sin\theta + \sin\alpha)$ appearing in the odd orders is the usual Fraunhofer double slit phase factor equaling half the phase difference in the single scattered waves arriving at an observation point P from corresponding parts of the two cylinders. The quantity $\gamma' = \gamma - 2Kb\sin\alpha$ in the even orders possesses the additional phase difference of the primary excitation of the two cylinders, agreeing with the above interpretation that the even orders of each cylinder arise from the first order of the remaining cylinder. Similarly the odd multiple orders of each cylinder possess only γ since these can be traced back to the cylinder's own first order of scattering which has been merely rescattered by the remaining cylinder.

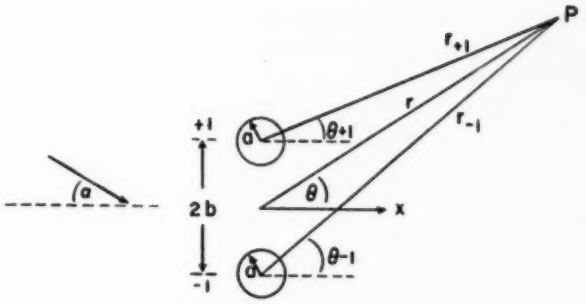


FIG. 2. Two cylinders.

Summing the geometrical progressions in (6) yields

$$\psi = \psi^1 + \sum_e \psi^m + \sum_o \psi^m = 2H_0\{C\cos\gamma + H(\Gamma_e + HC_r\Gamma_0)/[1-(HC_r)^2]\}, \quad (7)$$

provided that $|(HC_r)^2| < 1$ so that the present approximate representation is convergent. This expression con-

[5] R. W. Wood, Phil. Mag. 4, 396 (1902); 23, 310 (1912); Phys. Rev. 48, 928 (1935); J. Strong, Phys. Rev. 49, 291 (1936); L. R. Ingersoll, Astrophys. J. 51, 129 (1920), and Phys. Rev. 17, 493 (1921). Parallel anomalies have been investigated by C. H. Palmer but the results have not yet been published.

stitutes the far field scattered wave for two identical parallel cylinders with separation large compared to wavelength and radius, and separation small compared to the distance of observation. It is basically no more difficult to employ for numerical calculations than the solution for the single cylinder, the only other terms involved being simple trigonometric or algebraic functions.

For normal incidence or observation, α or $\theta=0$, we have $\Gamma_e=\Gamma_0\equiv\Gamma$ and (7) reduces to

$$\psi=2H_0[C\cos\gamma+\Gamma H/(1-HC_\pi)], \tag{8}$$

where $\gamma=Kb\sin\theta$ for $\alpha=0$. For either the forward or back scattered wave, $\theta=0$ or $\theta=\pi$, respectively, Γ reduces to $[\sum\epsilon_nA_n\cos(n\pi/2)]^2$, while $C\cos\gamma$ yields $\sum\epsilon_nA_n\equiv C^0$ for the former and C_π for the latter. Numerical calculations are particularly simple for these values of α and θ and attempts to verify these expressions should be undertaken.

For these cases, $\sin\alpha=0$, the primary excitations of the cylinders are in phase, as are, as a consequence, their excitations in any particular order of scattering; thus the waves leaving the two cylinders are initially in phase in each order of scattering. Furthermore, since $\theta=0$ or π, we are observing along the "central maximum" of $\cos\gamma$ so that the waves of each order of scattering arrive in phase at P. Hence from elementary physical considerations we expect that the effects of multiple scattering would be greatest, and yield an increase over the scattered intensity predicted by single scattering theory, for those wavelengths for which all orders were in phase. Were we to take into account only the phase factors arising from the number of times each multiply scattered wave (or rather its corresponding excitation) traversed the spacing $2b$, we would expect the greatest effects to occur for the wavelengths $\lambda=2b/p$, $2Kb=2p\pi$, which are integral fractions of the spacing. It is clear from our expressions, however, that all orders are not in phase for these wavelengths; they still differ by the phase factors arising from the complex scattering coefficients of the single cylinder, as well as by the multiples of $-i\pi/4$ which are a consequence of the cylindrical character of the waves. Thus the multiply scattered orders differ in phase from the first order by $\exp i[-arg(C)+arg(\Gamma)+(m-1)(2Kb-\pi/4)+(m-2)\times arg(C_\pi)]\equiv\exp(i\beta_m)$, and from each other by $\exp i[2Kb-\pi/4+arg(C_\pi)]\equiv\exp(i\beta)$. For the special cases where β_m reduces to $(m-1)\beta$, we can readily determine the "resonance" wavelengths for which all orders are in phase and reinforce the first from the requirement that $\beta=1$; similarly from $\beta=-1$ we can determine the wavelengths for which successive orders are out of phase and partially annul the first.

Similarly for arbitrary angles of incidence we would require the conditions $2Kb=2p\pi$ and $2Kb\sin\alpha=2p'\pi$, where p and p' are both integral or both half odd integral, or $2Kb(1\pm\sin\alpha)=2p_\pm\pi$, where p_+ and p_- are integral, to eliminate the phase differences arising from the primary excitation and the spacing. Although the orders are not in phase for these values, they simplify the general expression (7) considerably. Thus for $\gamma=q\pi$ (the "principal maxima" of $\cos\gamma$) and for these conditions, (7) reduces to

$$\psi=2H_0(-1)^q[C+hG_e/(1-hC_\pi)+hG_o/(1+hC_\pi)], \tag{9}$$

where $h=(i\pi Kb)^{-1}$ and where $G=\sum\epsilon_nA_n\cos n(\theta+\pi/2)\times\sum\epsilon_{n'}A_{n'}\cos n'(\alpha-\pi/2)$, the subscripts e and o indicating that n and n' are both even or both odd. It can be seen that the multiply scattered orders have reduced to two components; the even component containing G_e for which successive orders differ in phase by $arg(C_\pi)-\pi/4$, and the odd containing G_0 for which successive orders differ by $arg(C_\pi)+\pi-\pi/4$. Hence for the wavelengths satisfying $2Kb(1\pm\sin\alpha)+arg(C_\pi)-\pi/4=2p_\pm\pi$, all terms of the even component are in phase while those of the odd are 180° out of phase, and conversely for $2p_\pm\pi$ replaced by $(2p_\pm+1)\pi$. The two components, however, are out of phase with each other as well as with the first order, and successive orders of scattering of the configuration are neither in nor out of phase, nor is it clear that resonance wavelengths exist for the general case.

For the limiting case of $Ka\to0$, however, and for certain boundary conditions, we will be able to determine by inspection the wavelengths for which the "maximal effects" of multiple scattering occur, or for which the departures from single scattering theory are most pronounced.

Mode Approximations, Ka→0

In this section we will obtain approximate solutions for the case of cylinders whose radii are very small compared to wavelength. For the single cylinder and $Ka\equiv X\to0$ the series representation for the scattered wave is very rapidly convergent so that retaining only the first three terms of C, or $A_0+2A_1\cos(\theta+\alpha)+2A_2\cos2(\theta+\alpha)$, yields a solution correct to X^4. For two (or more) cylinders there are essentially two procedures we can follow; either we expand (7) as a power series in X and retain only all terms to a specified power of X (so that we neglect all but the first few orders of scattering), or else we retain all orders of scattering and neglect all but the leading terms in X. To illustrate the first procedure, we obtain a solution correct to X^4 by retaining only A_0, A_1, and A_2 in ψ^1 and A_0^2, A_0A_1, A_1^2 in ψ^2; all the remaining individual terms of the ψ^m are at least as small as X^6 [and also smaller because of $(Kb)^{-(m-1)/2}$]. Because of the product form of the ψ^m, however, there is a rapid increase of the number of terms containing each power of X (particularly, if there are many cylinders) so that the sum of all terms containing X to powers greater than 4 may be appreciable and of the same order of magnitude as the terms retained.

The alternate procedure, say retaining only all terms containing A_0 and A_1 in all orders of scattering, is based

on the fact that the A_n correspond to the modes of oscillation of one cylinder. Thus if we assume, for example, that each cylinder can oscillate only in the "A_0 mode," or that each scatters isotropically regardless of the wave form of the incident radiation, we would retain only the terms containing A_0 in all orders of scattering.[6] While this assumption seems physically plausible, it is apparent from our expressions that by its use we introduce discrepancies in regard to the powers of X retained (in addition to the discrepancies in powers of (Kb) introduced previously in using the asymptotic form of (4)). The virtue of this procedure, which we employ primarily for heuristic purposes, is that all orders of scattering are retained, albeit only the largest term of each, to yield a physically plausible, explicit solution which can be analyzed without resorting to numerical computations. In this regard we note that this procedure when applied to the problem of the grating[7] yields good agreement with the experimental data of reference 5.

We will consider the three "modes," A_0, A_1, and A_0+A_1.

"A_0 Mode" Isotropic Scatterers

For an incident plane wave polarized parallel to a cylinder's axis, or for a cylinder of zero impedance in acoustics, and $Ka\to 0$, the largest coefficient is $A_0 = iB_0e^{i\varphi}$, where B_0 and φ are real and positive. Thus for a nonferromagnetic conducting dielectric $B_0=\pi X^2 \times[(k_e-1)^2+(\sigma/\epsilon\omega)^2]^{\frac{1}{2}}/4$, $\varphi=\tan^{-1}[\sigma/\omega(\epsilon'-\epsilon)]$; for a perfect dielectric $B_0=\pi X^2(k_e-1)/4$, $\varphi=0$; and for a perfect conductor or a cylinder of zero impedance $B_0=\pi/2\ln(2/cX)$, $\varphi=\pi$. (The notation is for mks units as in Stratton,[8] $k_e=\epsilon'/\epsilon$, and $c=1.781$.) For the first two cases we write $B_0=\pi X'^2/4$ for brevity.

Invoking the mode hypothesis yields

$$\psi=\sum\psi^m=\psi^1(1+A_0H\cos\gamma'/\cos\gamma)/(1-A_0^2H^2), \quad (10)$$

where $\psi^1=2H_0A_0\cos\gamma$. The physical interpretation of this solution is that each cylinder is excited by a plane wave and by a superposition of symmetrical cylindrical waves having a common origin, which in effect are plane on arrival because of $2Kb\gg 1$, $Ka\ll 1$. The cylinder responds to both of these excitations by scattering isotropically, or the cylinder can but radiate in the "A_0 mode."

An analytic criterion for the range of validity of the single scattering approximation is furnished by the requirement that $B\equiv|(\psi-\psi^1)/\psi|\ll 1$, or that the sum of the multiple orders of scattering be small compared to the total scattered wave. Thus when $\gamma=\gamma'$ we have $B=|A_0H|=|\psi^2/\psi^1|=B_0(\pi Kb)^{-\frac{1}{2}}$. The logarithmic case, $B=(\lambda/2b)^{\frac{1}{2}}/2\ln(2/cX)$, is the worst as far as the usefulness of single scattering theory is concerned. Thus for $X=0.5$, 0.05, and 0.005, the ratio of spacing to wavelength must be 144, 9, and 4, respectively, for the higher orders of scattering to contribute less than 5 percent. For the other case, $B=(\lambda/2b)^{\frac{1}{2}}(X'/2)^2$, so that even relatively large values such as $X'\approx 0.3$ and spacing of the order of magnitude of the wavelength yield $B\approx 5$ percent.

We note that this criterion could have been obtained simply and more directly from rather elementary physical considerations. For these ranges of the parameters, the solution of the single cylinder equals $H_0(2Kb)A_0$ on arrival at the second cylinder. The wave, the second cylinder scatters in response, is then proportional to this excitation factor, so that the smallness of the factor insures that the second-order wave scattered by the cylinder is negligible; similarly for the higher orders. Furthermore this same sort of approach could have been employed to write down the scattered wave for this problem immediately. Thus we generalize (10) by inspection and by means of these considerations to the case where the cylinders have different parameters to obtain

$$\psi=H_0[(A_0+A_0')\cos\gamma + 2A_0A_0'H\cos\gamma']/(1-A_0A_0'H^2), \quad (10')$$

which also includes the case of spherical scatterers if the appropriate coefficients and functions are substituted.

For the present case the discussion following (8) can be extended to determine the actual wavelengths for which the effects of multiple scattering are maximal. Rewriting (10) as

$$\psi=H_0A_0[e^{i\gamma}(1+Be^{i\beta-i2A}) + e^{-i\gamma}(1+Be^{i\beta+i2A})]/(1-B^2e^{i2\beta}), \quad (11)$$

where $\beta=2Kb+\varphi+\pi/4$ and $A=Kb\sin\alpha$, the maximal wavelengths can be determined by inspection. Thus for wavelengths satisfying

$$2Kb(1\pm\sin\alpha)+\varphi+\pi/4=2p_\pm\pi, \quad p_\pm=1,\ 2,\ 3\cdots, \quad (12)$$

(11) reduces to

$$\psi=2H_0A_0\cos\gamma/(1-B)=\psi^1(1+B+B^2\cdots), \quad (13)$$

all orders of scattering being in phase and reinforcing the first to yield intensity maxima. Similarly for wavelengths satisfying

$$2Kb(1\pm\sin\alpha)+\varphi+\pi/4=(2p_\pm+1)\pi, \quad (12')$$

we obtain

$$\psi=2H_0A_0\cos\gamma/(1+B)=\psi^1(1-B+B^2-B^3\cdots), \quad (13')$$

successive orders of scattering being 180° out of phase and partially annulling the first to decrease the intensity predicted by single scattering theory. Hence the "principle maxima" of the single scattered wave as determined by $\gamma=q\pi$ will be most intensified at angles of ob-

[6] Essentially this sort of assumption is tacit in most previous work on multiple scattering.

[7] V. Twersky, "Multiple Scattering of Radiation, Part II," Research Report EM-39, Math. Research Group, W.S.C., New York University (1951).

[8] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941).

servation given by $\sin\theta=(q-p')/(p-\varphi/2\pi-\frac{1}{8})$, where $p'=(p_+-p_-)/2$ and $p=(p_++p_-)/2$, and most reduced for the values obtained by replacing $\frac{1}{8}$ by $\frac{5}{8}$ in this relation. We also note that for these wavelengths multiple scattering has no effect on the absolute minima of the scattered intensity as, determined by $\cos\gamma$, since both ψ^1 and ψ vanish for $\gamma=q_0\pi/2$ (q_0 being odd). In general, however, only the odd orders of scattering vanish for this condition, and the further condition $2A=q\pi$ is required for the total scattered wave to vanish; the waves scattered by each cylinder are then completely out of phase with those from the other on arrival at P.

It can be seen that if the wavelengths satisfy only one of the conditions of either (12) or (12′) the effects of multiple scattering are also relatively strong. Thus for $\beta+2A=2p_+\pi$, (11) reduces to

$$\psi=H_0A_0\{e^{i\gamma}+(1+B)e^{-i\gamma}+(e^{i\gamma}+Be^{-i\gamma})B(1+B)e^{i2\beta}/(1-B^2e^{i2\beta})\}. \quad (14)$$

Hence for $\gamma=q\pi$ we obtain

$$\psi=H_0A_0(-1)^q\{2+B+B(1+B)^2e^{i2\beta}/(1-B^2e^{i2\beta})\}, \quad (15)$$

which we interpret as increasing the scattered intensity relative to its singly scattered value; the first order of scattering of both cylinders being in phase with the second order of -1, the second order of $+1$ being in phase with the third order of both and with the fourth order of -1, the fourth order of $+1$ being in phase with the fifth order of both and the sixth order of -1, etc.

Similarly for one of the conditions of (12′) there is a certain amount of annulment among the various orders of scattering. Thus for $\beta+2A=(2p_++1)\pi$, we obtain (14, 15) with B replaced by $-B$, and interpret the second of these as decreasing the scattered intensity relative to its single scattered value.

Plotting K *vs* α for the set of conditions of (12) yielding an increase, for both $\pm2A$, yields a set of straight lines forming a "network." Similarly for the set yielding a decrease, (12′), we obtain the same graph but the two networks are displaced from each other along the K-axis by $\pi/2b$. Physically the plot would correspond to a series of correlated "spectrograms" for different angles of incidence in which the lines of the first network correspond to bright bands, and the lines of the second to dark bands. We note that for the perfect conductor $B\propto\lambda^{\frac{1}{2}}/\ln(\lambda/c\pi a)$, so that the most pronounced departures from single scattering theory occur for the longer wavelengths, while for the dielectrics $B\propto\lambda^{-1}$, and the effects of the multiple scattering fall off rapidly with increasing wavelength.

"A_1 Mode"

For a plane wave polarized perpendicular to the axis of a dielectric cylinder or for a gaseous cylinder in acoustics, and $Ka\to0$, the largest coefficient is $A_1=iB_1e^{i\varphi}$. Thus for a nonferromagnetic conducting dielectric $B_1=(\pi X^2/4)\{[(k_e-1)^2+(\sigma/\epsilon\omega)^2]/[(k_e+1)^2+(\sigma/\epsilon\omega)^2]\}^{\frac{1}{2}}$, $\varphi=\tan^{-1}\{2\sigma/\epsilon\omega[k_e^2-1+(\sigma/\epsilon\omega)^2]\}$, while $B_1=\pi X^2(k_e-1)/4(k_e+1)$, $\varphi=0$, for a perfect dielectric (or for a gaseous cylinder if $k_e=\rho'/\rho$ is the ratio of the densities). Retaining only this coefficient in the complete solution yields

$$\psi=\psi_e^1+\psi_o^1(1+2A_1H\cos\gamma'/\cos\gamma)/[1-(2A_1H)^2]=\psi_e+\psi_o, \quad (16)$$

where $\psi^1=4H_0A_1(\cos\theta\cos\alpha-\sin\theta\sin\alpha)\cos\gamma=\psi_e^1+\psi_o^1$ (e and o indicating the even and odd components, respectively). It can be seen that to this order of approximation in Kb only the component symmetrical with respect to the plane of the configuration is effected by multiple scattering, the solution reducing to its single scattered value for either normal incidence or observation. Physically each cylinder is excited by a plane wave and by a superposition of cylindrical waves whose amplitude is a maximum along the line joining the cylinders' axes. The cylinder responds to each of these excitations by scattering a cylindrical wave having maximum amplitude along the direction of incidence (or that determined by the two axes), or the cylinders can but radiate in the "A_1 mode."

We note that ψ_o is formally identical with (10) so that most of the previous material is pertinent for the present case as well. For the maximal conditions of (12) and (12′) with the appropriate values of φ we obtain

$$\psi=\psi_e^1+\psi_o^1/(1\mp2B), \quad (17)$$

where $B=B_1(\pi Kb)^{-1}$. Hence for small angles of incidence or observation the intensity differs little from its single scattered value. As α and θ increase, however, the odd component becomes more effective; finally, for either grazing incidence or observation the expressions for the total scattered wave are formally identical with those discussed previously for isotropic scatterers. We see, therefore, that for large α or θ for a perfect dielectric, maximal effects of the same type and of the same order of magnitude (neglecting the $\cos\theta\cos\alpha$ factor) occur for both the perpendicular and parallel components, and at the same wavelengths. For the conducting dielectric the phase factors φ are no longer zero, and the maximal effects will not in general occur simultaneously for both components at the same wavelengths.

"A_0+A_1 Mode"

For a plane wave polarized perpendicular to the axis of a perfectly conducting (or ferromagnetic) cylinder, or for a rigid cylinder or one of large impedance in acoustics, and $Ka\to0$, the largest coefficients are $A_0=iB_0e^{i\varphi_0}$ and $A_1=iB_1e^{i\varphi_1}$. Retaining only these terms yields $\psi^1=2H_0[A_0+2A_1\cos(\theta+\alpha)]$ and

$$\psi=\psi^1+2H_0H\{(A_0^2-4A_1^2\sin\theta\sin\alpha)\cos\gamma'-i2A_0A_1(\sin\theta-\sin\alpha)\sin\gamma'+H(A_0-2A_1)\times[(A_0^2+4A_1^2\sin\theta\sin\alpha)\cos\gamma-i2A_0A_1(\sin\theta+\sin\alpha)]\}[1-H^2(A_0-2A_1)^2]^{-1}, \quad (18)$$

which is the sum of the A_0 and A_1 modes plus "coupling terms" in A_0A_1. If $2A=2p'\pi$ and $\gamma=q\pi$, this reduces (essentially as for (9)) to

$$\psi={}^{e}\psi^{1}(1+2HA_1)/[1-H(A_0-2A_1)] +{}^{1}\psi_{o}{}^{1}(1+A_0H)/[1+H(A_0-2A_1)]+{}^{1}\psi_{e}{}^{1}, \quad (19)$$

where ${}^{e}\psi^{1}$ is the first-order scattered wave containing only A_0, and similarly ${}^{1}\psi^{1}$ contains only A_1; e and o indicate even and odd as previously.

For normal incidence or observation ${}^{1}\psi_{o}{}^{1}$ vanishes and ψ is similar to the forms discussed previously, although complicated by the presence of both A_0 and A_1; the multiplier of ${}^{e}\psi^{1}$ is real for the conditions of (12, 12') with $\varphi=arg[A_0/(1+2HA_1)]$. For very large α and θ the character of the multiple scattering effects are determined primarily by the multiplier of ${}^{1}\psi_{o}{}^{1}$ which is real for the conditions of (12, 12') with $\varphi=arg[-2A_1/(1+A_0H)]$. In general, the two terms (corresponding to the even and odd components of (9)) "compete" to determine the character of the effects.

For a perfectly conducting or rigid cylinder, $A_0=-A_1=-i\pi X^2/4=iB_0e^{i\tau}$ and the multipliers in (19) reduce to $(1-2A_0H)/(1-3A_0H)$ and $(1+A_0H)/(1+3A_0H)$. Hence for the wavelengths of (12) and (12') we obtain

$$\psi=\psi^{1}\mp B'B[(1\pm 3B)^{-1}-4(1\mp 3B)^{-1}\sin\theta\sin\alpha], \quad (20)$$

where $\psi^{1}=B'[1-2\cos(\theta+\alpha)]$, $B'=(-1)^{q}X^2(\pi/2Kr)^{\frac{1}{2}}$, and $B=X^2(\pi/Kb)^{\frac{1}{2}}/4$; the top and bottom signs correspond respectively to the conditions of (12) and (12'). It can be seen that although for certain wavelengths $|\psi/\psi^{1}|^2$ may be greater than unity for normal incidence, it may be less than unity for larger angles, and conversely. Comparison of (20) with the result for parallel polarization for the perfect conductor indicates that for α small and θ large the bands are of the same type for the two cases, while the bands are of opposite type for α and θ large; the effects are also much more pronounced for the parallel component. The forward scattered wave, $\theta=-\alpha$, is $\psi=B'\{-1\mp B[(1\pm 3B)^{-1}+4(1\mp 3B)^{-1}\sin^2\alpha]\}$, while the backscattered wave, $\theta=\pi-\alpha$, is $\psi=B'\{3\mp B[(1\pm 3B)^{-1}-4(1\mp 3B)^{-1}\sin^2\alpha]\}$. Thus for small α and wavelengths satisfying (12), the intensity of the former is increased while that of the latter is decreased, and conversely for the wavelengths of (12'). For large α, however, both intensities are increased for the wavelengths of (12) and both decreased for those of (12').

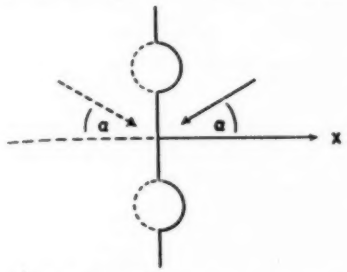


FIG. 3. Two semi-cylindrical bosses on a plane.

For simplicity, we have considered only the scattered wave. For comparison with experiment, however, one must also take into account the incident plane wave to derive the total intensities which are usually measured. For acoustics the total mean square pressure and radial component of the acoustic intensity to the order of r^{-1} are given by

$$P^2=(\rho^2\omega^2/2)[1+2f+|\psi|^2], \quad I=(K\omega\rho/2)\{\cos(\theta+\alpha) +[1+\cos(\theta+\alpha)]f+|\psi|^2\}, \quad (21)$$

where $f=Re\{\psi\exp[-iKr\cos(\theta+\alpha)]\}$ indicates the interference of the incident and scattered wave. For electromagnetics we specify the incident wave by $\mathbf{E}_i=\{E''[\sin(\theta+\alpha)\mathbf{r}_0+\cos(\theta+\alpha)\boldsymbol{\theta}_0]-E'\mathbf{k}_0\}\exp[iKr\times\cos(\theta+\alpha)]$, where E'' and E' are the polarization components perpendicular and parallel to the cylinders' axes. The total polarization intensity and radial Poynting flux to the order of r^{-1} are then

$$J=\mathbf{E}\cdot\mathbf{E}^*/2=\{|E''|^2[1+2\cos(\theta+\alpha)f''+|\psi''|^2] +|E'|^2[1+2f'+|\psi'|^2]\}/2,$$
$$S=(K/2\mu\omega)[|E''|^2\{\cos(\theta+\alpha)+[1+\cos(\theta+\alpha)]f'' +|\psi''|^2\}+|E'|^2\{\cos(\theta+\alpha) +[1+\cos(\theta+\alpha)]f'+|\psi'|^2\}], \quad (22)$$

the values of f'', $|\psi''|^2$ for the perpendicular component, and those for f', $|\psi'|^2$ for the parallel component being obtained from the derived ψ's with the appropriate values of the A's for the single cylinder. Since f is, in general, greater than $|\psi|^2$ for $Ka\ll 1$, its extrema determines those of the total radiation.

4. TWO BOSSES ON AN INFINITE PLANE

For the analogous problems of nonspecular reflection from two bosses on an infinite perfectly conducting plane, as in the right half of Fig. 3, the total reflected wave and its scattered component are obtained by the image method discussed in section 2 and indicated schematically in Fig. 3. Thus from the explicit form of ψ in (7) we have

$$\psi''=4H_0\{C_+\cos\gamma+H(\Gamma_e+HC_r\Gamma_0)/[1-(HC_r)^2]\}, \quad \psi'=4H_0C_-\cos\gamma, \quad (23)$$

where $\psi''=\psi(\theta)+\psi(\pi-\theta)$ and $\psi'=\psi(\theta)-\psi(\pi-\theta)$ are respectively the scattered components polarized perpendicular and parallel to the bosses, $C_+=\sum\epsilon_nA_n\times\cos n(\alpha+\pi/2)\cos n(\theta-\pi/2)$, and $C_-=-2\sum A_n\sin n(\alpha+\pi/2)\sin n(\theta-\pi/2)$. Similarly for the mode approximations for dielectric bosses we have $\psi''=2\psi_0$ with ψ_0 as in (16), while $\psi'=0$; ψ'' is also the scattered wave for gaseous bosses on a rigid plane. For perfectly conducting bosses, $\psi''=2\psi_0$ where ψ_0 is ψ of (18) without the term in $\cos\theta\cos\alpha$, while $\psi'=8A_1'H_0\cos\theta\cos\alpha\cos\gamma$ (where A_1' equalling A_0 as for the perpendicular component, was previously neglected compared to the logarithmic coefficient); ψ'' is also the solution for a rigid surface.

As a consequence of the symmetry of the multiply scattered orders with respect to the plane for $Kb\gg 1$,

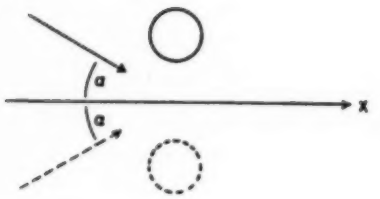


FIG. 4. Single cylinder parallel to a plane.

departures from single scattering theory are indicated only for the perpendicular component to lowest order in Kb. Proceeding to the next order of magnitude in Kb for the parallel component, however, we obtain from (3)

$$\psi' = 8A_0H_0(r)\cos\theta\cos\alpha[\cos\gamma + A_0(H_0+H_2)\cos\gamma']/[1-A_0^2(H_0+H_2)^2], \quad (24)$$

the effects being much less for this component, since $H_0+H_2=2H_1/Kb\approx -i2H_0/Kb$. For the conditions $\beta' = 2Kb\pm 2A-\pi/4=2p_\pm\pi$, and $\beta'=(2p_\pm+1)\pi$, corresponding respectively to (12) and (12′), (24) reduces to

$$\psi' = 4B'(1\pm 2B/Kb)^{-1}\cos\theta\cos\alpha, \quad (24')$$

where B' and B are as in (20). We note that not only are the maximal effects much less for this case than for two cylinders and parallel polarization discussed in the previous section, both because of the factor Kb and the fact that the previous A_0 was logarithmic in (Ka), but also that the peaks of the effects fall at shorter wavelengths for the present case. We also note that the present effects decrease with increasing λ while for two cylinders they increased.

For an incident wave specified by $\mathbf{E}_i(\pi-\theta)=\{-E''\times[\sin(\theta-\alpha)\mathbf{r}_0+\cos(\theta-\alpha)\boldsymbol{\theta}_0]+E'\mathbf{k}_0\}e^{-iKr\cos(\theta-\alpha)}$ the total reflected intensities are given by (22) with the present values of the ψ's, the wave $E_i=E_i(\theta)$ stated previously for the two cylinders being the specularly reflected component for the present problem.

5. CYLINDER AND INFINITE PLANE

The solution of the scattering problem for two cylinders enables us to treat the scattering reflection problem of a cylinder parallel to an infinite nonabsorbing plane as in the top half of Fig. 4; we will employ the image method indicated in the figure. For the component polarized perpendicular to the cylinder's axis (or for a rigid plane), the incident wave for the image problem is $\psi_i(\theta)+\psi_i(-\theta)$, where $\psi_i(-\theta)$ is the original ψ_i with θ replaced by $-\theta$; the total scattered wave is $\psi^+=\psi(\theta)+\psi(-\theta)$; while the sum of the scattered and reflected waves for the cylinder and plane problem is $\psi_r{}^+=\psi^+ +\psi_i(-\theta)$. Similarly for the component polarized parallel to the cylinder's axis the total incident wave is $\psi_i(\theta)-\psi_i(-\theta)$, the scattered wave is $\psi^-=\psi(\theta)-\psi(-\theta)$, while $\psi_r{}^-=\psi^- -\psi_i(-\theta)$. Thus from (7) we obtain

$$\psi^\pm = 2H_0[C(\alpha+\theta)\cos\gamma\pm C(\alpha-\theta)\cos\gamma' + H(\Gamma_s\pm\Gamma_0)/(1\mp HC_s)]. \quad (25)$$

Both scattered components therefore show maximal effects of the type previously discussed.

Subject to the mode hypothesis we obtain for A_0, A_1, and A_0+A_1, respectively,

$$\psi^- = -4H_0A_0\sin T\sin A/(1+HA_0),$$
$$\psi_1{}^+ = 8H_0A_1[\cos T\cos A\cos\theta\cos\alpha + \sin T\sin A\sin\theta\sin\alpha/(1+2HA_1)],$$
$$\psi^+ = 4H_0[\cos T\cos A(A_0+2A_1\cos\theta\cos\alpha) + 2A_1\sin T\sin A\sin\theta\sin\alpha]+4H_0H[1-H(A_0-2A_1)]^{-1}[A_0{}^2\cos T\cos A-4A_1{}^2\sin T\sin A \times\sin\theta\sin\alpha-i2A_0A_1(\sin T\cos A\sin\theta + \cos T\sin A\sin\alpha)], \quad (26)$$

where $T=Kb\sin\theta$ and $A=Kb\sin\alpha$.

The parallel polarization component, ψ^-, vanishes when α or $\theta=0$, or for grazing incidence or emergence, and has its "principal maxima" when T and A equal an odd multiple of $\pi/2$. The maximal effects of multiple scattering are determined by the conditions of (12) and (12′), the former yielding the minima and the latter the maxima for the present case. Similarly for the perpendicular component for the dielectric cases, $\psi_1{}^+$, the odd term behaves essentially as ψ^- for α or θ large (in the vicinity of the surface normal), while the even component is free of multiple scattering and has its maxima for T and A equalling an integral multiple of π. For ψ^+ with two coefficients, as for the perfect conductor or rigid cylinder, etc., the situation is more complicated. It is simpler than for two cylinders or two bosses, however, inasmuch as the odd and even components have identical denominators; hence the previous competition for the character of the bands is absent.

The writer is grateful to Professor Joseph B. Keller for his aid and to Professor Fritz Reiche and Dr. Jerry Shmoys for their interest.

APPENDIX

The electromagnetic field is obtained from the Hertz potentials by means of

$$\mathbf{E}=\frac{iE''}{Kr}\frac{\partial\Psi''}{\partial\theta}\mathbf{r}_0-iE''\frac{\partial\Psi''}{\partial(Kr)}\boldsymbol{\theta}_0-E'\Psi'\mathbf{k}_0,$$
$$\mathbf{H}=\eta\left\{\frac{iE'}{Kr}\frac{\partial\Psi'}{\partial\theta}\mathbf{r}_0-iE'\frac{\partial\Psi'}{\partial(Kr)}\boldsymbol{\theta}_0+E''\Psi''\mathbf{k}_0\right\},$$

where $\eta=K/\mu\omega=(\epsilon/\mu)^{\frac{1}{2}}$ in mks units, E'' and E' are the polarization amplitudes perpendicular and parallel to the cylinders' axes, and $\Psi''{}'=\psi_i+\psi''{}'$. The wave functions ψ'' and ψ' are as in (1), (3), etc., and differ only in the values of the ${}^sA_n$ they contain, the ${}^sA_n$ being determined from the usual boundary conditions for the single cylinder. For the far field scattered wave the above reduce to

$$\mathbf{E}^s=E''\psi''\boldsymbol{\theta}_0-E'\psi'\mathbf{k}_0, \quad \mathbf{H}^s=\eta[E'\psi'\boldsymbol{\theta}_0+E''\psi''\mathbf{k}_0].$$

The acoustic field, or particle velocity and excess pressure, is obtained by means of $\mathbf{V}=\nabla\Psi$, $p=i\omega\rho\Psi$; ρ is the density of the undisturbed medium.

Focusing in a Beta-Ray Spectrometer with Tilted Sources*

C. M. FOWLER AND P. DOMOTOR†
Physics Department, Kansas State College, Manhattan, Kansas
(Received October 11, 1951)

Formulas are given which express the theoretical line shape obtained in a semicircular beta-ray spectrometer. The sources are taken to be plane, uniform, and isotropic. Line shapes of a monoenergetic electron group are studied as a function of the angle between source and focal planes. The treatment given applies only to the electrons emerging perpendicular to the magnetic field. However, conditions simulating the requirements for qualitative applicability of the formulas may be obtained practically by slitting out electrons having appreciable momentum components along the field. Two figures are given which illustrate line shapes for various spectrometer geometries.

A KNOWLEDGE of the line shape formed by a monoenergetic electron group upon the focal plane of a beta-ray spectrometer is of importance in investigations dealing with quantitative electron radiation studies. It is also useful to serve as a guide for adjusting the spectrometer slit and source arrangements for either high resolution or large intensities. Several writers[1] have discussed phases of this problem, and some of these have considered, in a qualitative manner, aspects of the problem treated in this note: the effect of tilting the electron source at an angle with the focal plane.

The exact treatment given below assumes an isotropically emitting plane source and considers effectively a "two-dimensional" spectrometer. In practice this latter condition can be achieved by a fairly stringent slitting in the direction of the field. Attention is focused upon the line shape resulting from a monoenergetic group. Since the particles having appreciable momentum along the field are slitted out, all focused electrons will have a common radius of curvature, R, governed by the field intensity and electron energy.

Figure 1, with the field perpendicular to the paper, shows a plane section of the source, slits, and focal plane. The source of length $2a$ is inclined at angle α with respect to the focal plane, and its center is distant d from this plane. The trajectory of a typical particle is described by the large arc, center 0. The source coordinates of the particle are given by x, the source position, and θ, the initial ejection angle from the source normal. On the focal plane the coordinates are y, the perpendicular distance from d, and the angle φ.

Defining A by Eq. (1), the transformation equations (2) readily follow

$$A=(d\cos\alpha-y\sin\alpha)/R, \tag{1}$$

$$\sin\theta+\sin\varphi=A$$
$$x=y\cos\alpha+d\sin\alpha-R[\cos\theta+(1-(A-\sin\theta)^2)^{\frac{1}{2}}]. \tag{2}$$

If I_0 is the total disintegration rate of the source, then the radiation from the source element dx into angle $d\theta$ is simply

$$I(x,\theta)dxd\theta=(I_0/4\pi a)dxd\theta. \tag{3}$$

To evaluate the line density as a function of y it is necessary to integrate over the angle φ at the point y. Since φ, θ are in one-one correspondence for a given y, it is convenient to integrate over θ instead. In this case,

$$I(\theta,y)d\theta dy=I(\theta,x)J\begin{pmatrix}\theta,x\\ \theta,y\end{pmatrix}d\theta dy=\frac{\partial x}{\partial y}I(\theta,x)d\theta dy. \tag{4}$$

A simple calculation shows that

$$\frac{\partial x}{\partial y}=\cos\alpha+\frac{\sin\alpha(\sin\theta-A)}{[1-(A-\sin\theta)^2]^{\frac{1}{2}}}. \tag{5}$$

Thus integrating over θ, the line density at y becomes

$$I(y)dy=\frac{I_0dy}{4\pi a}\int_{\theta_1}^{\theta_2}d\theta\left\{\cos\alpha+\frac{\sin\alpha(\sin\theta-A)}{[1-(A-\sin\theta)^2]^{\frac{1}{2}}}\right\}. \tag{6}$$

The reduction to standard form of this elliptic integral requires a little care, in that the parameter A becomes negative when R is large enough that points y lie beyond the intersection of the extended source line with the focal plane. This situation only arises for positive angles of α. To conserve space the reduction is carried out only for $A>0$, but the final forms will be given for both cases.

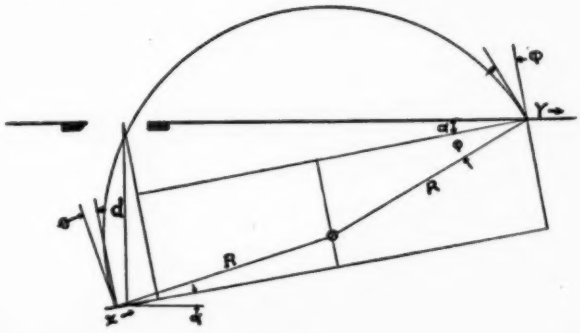


FIG. 1.

* Sponsored in part by the AEC.
† Now at the University of Illinois, Department of Physics.
[1] W. A. Wooster, Proc. Roy. Soc. (London) A114, 729 (1927); K. T. Li, Proc. Cambridge Phil. Soc. 33, 164 (1937); A Tyler and J. Lawson, thesis, University of Michigan (1939); K. Siegbahn, Arkiv Mat, Ostr. och Fysik, 30 (1944); G. E. Owen, Rev. Sci. Instr. 20, 916 (1949); Fowler, Shreffler, and Cork, Rev. Sci. Instr. 20, 966 (1949).

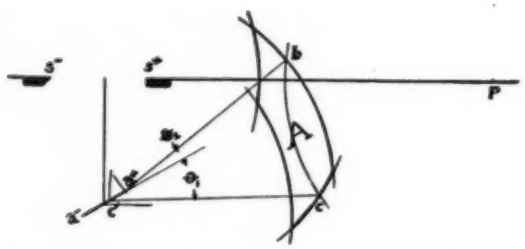


FIG. 2. Graphical construction for determination of angle limits.

Using the transformation of Eq. (7) and the terminology of Eq. (8) the intensity formula transforms into Eq. (9)

$$\sin^2\delta=\frac{1}{n}\left(\frac{1+\sin\theta-A}{1+\sin\theta}\right);\quad A>0. \tag{7}$$

$$k^2=1-\tfrac{1}{4}A^2;\quad n=1-\tfrac{1}{2}|A|. \tag{8}$$

$$I(y)=\frac{I_0}{4\pi a}\Big[\cos\alpha(\theta_2-\theta_1)+\sin\alpha\Big\{A\int_{\delta_1}^{\delta_2}\frac{d\delta}{(1-n\sin^2\delta)(1-k^2\sin^2\delta)^{\frac{1}{2}}}-(A+1)\int_{\delta_1}^{\delta_2}\frac{d\delta}{(1-k^2\sin^2\delta)^{\frac{1}{2}}}\Big\}\Big]. \tag{9}$$

Following the defining equations for the Jacobian zeta- and eta-functions[2] the elliptic integral of the third kind becomes

$$\int_0^{\delta}\frac{d\delta}{(1-n\sin^2\delta)(1-k^2\sin^2\delta)^{\frac{1}{2}}}=u+\frac{snv}{cnvdnv}\Big[uZnv+\tfrac{1}{2}\ln\frac{\theta(v-u)}{\theta(v+u)}\Big], \tag{10}$$

where u is the elliptic integral of the first kind, Eq. (11), and v is defined by Eq. (12)

$$u=\int_0^{\delta}\frac{d\delta}{(1-k^2\sin^2\delta)^{\frac{1}{2}}}. \tag{11}$$

$$k^2sn^2v=n. \tag{12}$$

For this case as follows from Eq. (8) and the definitions of the elliptic functions,

$$(Asnv)/(cnvdnv)=2. \tag{13}$$

[2] Smithsonian Misc. Coll., Vol. 74, No. 1, Pub. 2672.

Combining Eqs. (9) to (13), the final reduction becomes

$$I(y)=\frac{I_0}{4\pi a}\Big[\cos\alpha(\theta_2-\theta_1)+\sin\alpha\Big\{(2Znv-1)(u_2-u_1)+\ln\frac{\theta(v-u_2)}{\theta(v-u_1)}\frac{\theta(v+u_1)}{\theta(v+u_2)}\Big\}\Big]. \tag{14}$$

For the case of negative A, with the same definitions for k and n, the suitable transformation corresponding to Eq. (7) is

$$\sin^2\delta=1/n[(1+\sin\theta)/(1+\sin\theta-A)];\quad A<0. \tag{15}$$

The formula for the line intensity is the same as that of Eq. (14) except the factor $(2Znv-1)$ is replaced by $(2Znv-A-1)$.

The determination of the θ-limits and thus the δ and $u(\delta)$ limits according to Eqs. (7) and (11) can be obtained graphically with sufficient accuracy, as shown in Fig. 2. Here both sample and slit widths have been exaggerated for clarity of construction. The intersections of the four circles, all with radius R centered at a^+, a^-, s^+, and s^-, define an area A, which contains the centers of all the particle trajectory circles which can originate from the source and pass through the slits. To determine the θ-limits corresponding to an arbitrary point $y=P$, a circle of radius R is constructed with P as center. The intersections b and c of this circle with the periphery of the area A, determine the limiting angles θ_1 and θ_2. Since the point b is an intersection with the a^+ circle, θ_2 is measured from the positive source limit a^+. On the other hand, the positive slit s^+ determines θ_1, which is then measured from the intersection c' of the source and a circle, radius R, with c as center. For certain conditions, the area A may consist of two portions. One of these portions corresponds to particles emitted from the sample side of the source. For the other area, the particles are emitted in the opposite direction. When this situation arises, it is necessary to specify that the source backing be permeable or impermeable to the particles, or to assign some particle absorption figure. Figure 3 shows such a case, where the trajectories whose centers lie in the lower area are ejected from one side of the source while those whose centers are in the upper area are emitted from the other side of the source. As most sources are deposited on only one side of the backing, it is evident that under the arrangement of this figure some focused emission

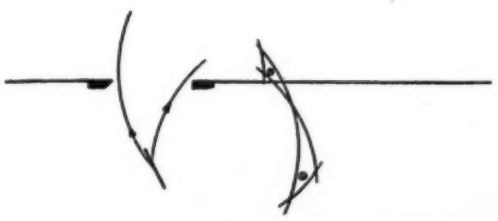


FIG. 3.

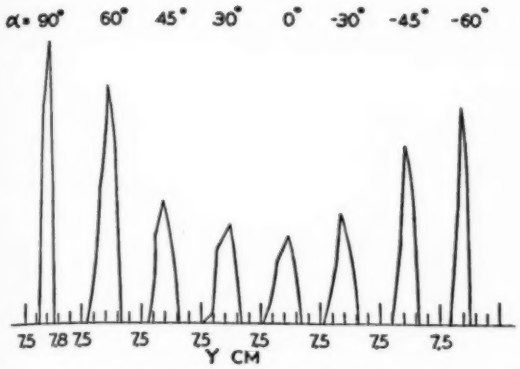


FIG. 4. Variation of line shape with α. Trajectory radius 4 cm, slit width 5 mm, (slit symmetrically spaced over source), a=1 mm, d=2 cm.

would have to pass through the backing. In practice such arrangements of the spectrometer geometry should be avoided.

Figures 4 and 5[3] present typical line shapes. In those instances where a portion of the electrons penetrate the backing, complete transmission is assumed. Figure 4 shows the continuous improvement in resolution and peak intensity as α departs from zero. This behavior agrees with that predicted qualitatively by earlier writers but is free of the divergences encountered there at $\alpha=90°$. Figure 5 compares the cases of $\alpha=0$ and $\alpha=90°$ for larger radii. The spectrometer geometry selected is typical of that used in variable field instruments. For these spectrometers the radii ordinarily are large enough to permit use of the following approximation formula when $\alpha=90°$. The data of Fig. 5 were so computed.

[3] Suitable tables of the eta, zeta, and other elliptic functions may be found in reference 2.

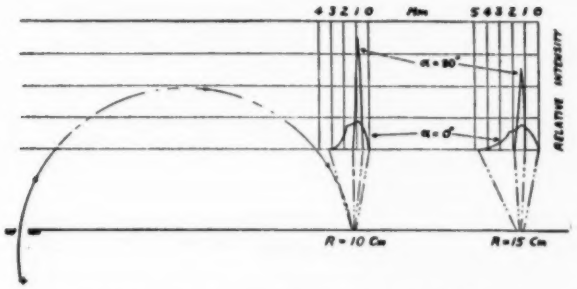


FIG. 5. Comparison of line shapes for $\alpha=0$, 90°, for radii 10 and 15 cm, slit width 4 mm (slit symmetrically spaced over source), a=1 mm, d=3 cm.

$$I(y)\cong\frac{I_0}{8\pi a}\left[\sin^{-1}\left(\frac{-y-2R\sin\theta_1}{2R-y}\right)-\sin^{-1}\left(\frac{-y-2R\sin\theta_2}{2R-y}\right)\right]. \quad (16)$$

This equation is derived from Eq. (6) directly upon replacing the variable θ by $w=A+1-\sin\theta$. For large R, the quantity $A=-y/R$ is nearly minus two, while $\sin\theta$ varies in the neighborhood of minus one. Thus w is small and only the two lowest powers of w have been retained.

Studies in Newtonian Flow. IV. Viscosity *vs* Molecular Weight in Liquids; Viscosity *vs* Concentration in Polymer Solutions

ARTHUR K. DOOLITTLE
Carbide and Carbon Chemicals Company, A Division of Union Carbide and Carbon Corporation, South Charleston, West Virginia

(Received October 15, 1951)

This is the final installment in the current series entitled "Studies in Newtonian Flow." In this paper, the molecular weight—free-space function applicable to *n*-paraffins $m=100$ through $m=240$ at various temperatures is further generalized to apply to homologous series of all Newtonian liquids at constant temperature. The resulting molecular weight function thereby loses some fidelity, and the slopes and intercepts no longer have physical significance. The generalized logarithmic decrement equation nevertheless reproduces viscosity data over extended ranges of molecular weight (at constant temperature) far better than expressions previously proposed. The equation is extended to apply to mixtures of liquids and finally to polymer solutions.

PREVIOUS papers of this series have discussed the dependence of the viscosity of liquids on temperature, on free-space, and on molecular weight and free-space in homologous series of pure liquids.[1] The studies on which the first three papers were based were made on members of the homologous series of *n*-alkanes. The original measurements were presented earlier in a paper[2] published elsewhere. We will in this paper refer again to the *n*-paraffin data, but will also introduce other viscosity data, some of which was obtained by us and some by others. Although all of our newly presented data are the result of careful measurement and are generally accurate to within 1 percent, we by no means claim the accuracy in the latter measurements that was sought for and obtained in the case of our *n*-alkane study.

GENERALIZED APPROXIMATE MOLECULAR WEIGHT FUNCTION

Unfortunately, the free-space concept, advanced in prior papers of this series, does not lend itself readily to the consideration of a variety of chemical families because of difficulties in accurately determining the magnitude of the free-space. A unique method of extrapolating density data to determine the specific volumes at absolute zero was employed in the case of the *n*-alkanes, but application of this method to other homologous series would be complicated not only by lack of data over a sufficient temperature range, but also by uncertainties in regard to the influence of association upon the final result. Therefore, for the purpose of approximate relations at least, it is desirable to express the viscosity directly as a function of molecular weight (at constant temperature) rather than as a more exact function of free-space which itself depends on molecular weight as well as on temperature. Such a generalized approximate function will now be developed.

The relatively accurate molecular weight—free-space equation developed in the previous paper is

$$\ln\eta = C(v_0/v_f)\exp(-K/m^{1/20}) + \ln A, \qquad (1)$$

where η = coefficient of viscosity, m = molecular weight, v_0/v_f = reciprocal relative free-space, C, K, A = constants. This equation represents the viscosity when both molecular weight and temperature vary. In the subsequent development we shall consider the case where the temperature is fixed and only the molecular weight is varied.

The right-hand side of Eq. (1) consists of a variable term and a constant term. Therefore, if $\ln\eta$ is plotted *vs* the variable, $\exp(-K/m^{1/20})(v_0/v_f)$, a straight line results with a slope C and intercept $\ln A$. Our object, therefore, should be to replace the variable of Eq. (1) with a variable that depends on molecular weight alone. Actually we do not make an equivalent replacement of $\exp(-K/m^{1/20})(v_0/v_f)_t$ by a function of m, but instead, we take advantage of a characteristic of logarithmic decrement functions that permits the curvature to be

TABLE I. *n*-Paraffin hydrocarbons. Viscosities at 100°C calculated by:
$\ln\eta_{100} = 2401\exp(-10.239/m^{1/20})\cdot(v_0/v_f) - 7.6500$ (3); $\ln\eta_{100} = 30.74\exp(-8.0/m^{1/4}) - 8.618$ (4).

m	$\exp(-10.239/m^{1/20})\cdot(v_0/v_f)$	$\exp(-8.0/m^{1/4})$	η_{obs}	$\eta_{calc(3)}$	% dev.	$\eta_{calc(4)}$	% dev.
100.20	0.0006137	0.07977	0.002099	0.002077	−1.0	0.002100	0.0
114.22	0.0007001	0.08654	0.002554	0.002556	+0.1	0.002586	+1.3
128.25	0.0007803	0.09280	0.003087	0.003099	+0.4	0.003135	+1.6
156.30	0.0009259	0.10408	0.004384	0.004397	+0.3	0.004434	+1.1
184.35	0.0010539	0.11405	0.006001	0.005978	−0.4	0.006024	+0.4
240.46	0.0012755	0.13114	0.01018	0.01018	0.0	0.01019	+0.1

[1] A. K. Doolittle, J. Appl. Phys. **22**, 1031, 1471 (1951); **23**, 236 (1952).
[2] A. K. Doolittle and R. H. Peterson, J. Am. Chem. Soc. **73**, 2145 (1951).

TABLE II. *n*-Paraffin hydrocarbons. Dependence of viscosity on molecular weight at 100°C.[a]
$\ln\eta_{calc} = 25.69 \exp(-6.7/m^{1/4}) - 9.277$.

m	$m^{1/2}$	$1/m^{1/4}$	$6.7/m^{1/4}$	$\exp(-6.7/m^{1/4})$	$25.69 \cdot \exp(-6.7/m^{1/4})$	η_{calc}	η_{obs}	% dev. (author)	% dev.[b] (Flory)
100.20	10.100	0.3161	2.1177	0.1203	3.090	0.002056	0.002099	−2.0	+0.6
114.22	10.687	0.3059	2.0495	0.1288	3.309	0.002557	0.002554	+0.1	+0.5
128.25	11.325	0.2972	1.9910	0.1366	3.508	0.003121	0.003087	+1.1	0.0
156.30	12.502	0.2828	1.8949	0.1503	3.862	0.004446	0.004384	+1.4	−1.1
184.35	13.578	0.2714	1.8183	0.1623	4.169	0.006046	0.006001	+0.8	−1.4
240.46	15.507	0.2539	1.7014	0.1824	4.686	0.01014	0.01018	−0.4	+1.5
394.74	19.868	0.2244	1.5031	0.2224	5.714	0.02834	0.02872	−1.3	+26.7
506.95	22.516	0.2108	1.4120	0.2437	6.259	0.04888	0.04884	+0.1	+60.1
899.68	29.995	0.1826	1.2234	0.2942	7.558	0.1791	0.1785[c]	+0.3	+280.

[a] This study makes use of all the observed values regardless of proximity to freezing points.
[b] From $\ln\eta_{calc} = 0.2888m^{1/2} - 9.052$. This is the least square equation for data over range $m = 100.20$ to $m = 240.46$.
[c] By extrapolation from 110.1°C.

controlled to a limited extent by choice of the exponential constant.

A plot of $\ln\eta$ *vs* $(v_0/v_f)_t$ gives a curve that is slightly concave upward. The curvature of $\ln\eta$ *vs* $\exp(-K/m^{1/20})$ is subject to control by the value assigned to the constant, K. Too low a value of K will make the curve concave upward, too high a value, concave downward. Consequently, a proper choice of K will rectify the product of $\exp(-K/m^{1/20})(v_0/v_f)_t$. It would be possible to choose a value of K, say K', that would rectify $\ln\eta$ *vs* $\exp(-K'/m^{1/20})$. This would, of course alter the value of the constants C and $\ln A$ of Eq. (1). The adaptability of the logarithmic decrement function to control of the curvature in the manner described depends to some extent, however, on the root or power of the variable. At least the 1/3 or 1/4 power seems to be required although higher roots appear satisfactory. Since the 1/20 power is rather unwieldy we have chosen the 1/4 power for use in the subsequent development. Therefore, to rectify viscosity data in terms of molecular weight only, we propose the following generalized approximate equation,

$$\ln\eta_t \cong \alpha \exp(-\beta/m^{1/4}) + \gamma, \qquad (2)$$

in place of the more limited but more accurate Eq. (1). It should be understood that the α and γ of Eq. (2) bear no relation to the C and $\ln A$ of Eq. (1). In the subsequent use of Eq. (2), the subscript, t, and the approximation symbol, $\sim$, will be dropped. It will be understood that the equation is to be used with data taken at constant temperature and that the relation represents an approximation rather than an exact equation.

Assignment of the appropriate numerical values to the constants of Eqs. (1) and (2) that apply to the *n*-paraffins over the range from $m = 100$ through $m = 240$ at 100°C, gives:

$$\ln\eta_{100} = 2401(v_0/v_f) \exp(-10.239/m^{1/20}) - 7.6500, \quad (3)$$

and

$$\ln\eta_{100} = 30.74 \exp(-8.0/m^{1/4}) - 8.618, \qquad (4)$$

Figure 1 from Table I shows the relative position of these lines. The degree of conformance may be read from the table.

Extension of the molecular weight range to $m = 900$ introduces some curvature when the data are plotted according to Eq. (4) and requires a change in β from

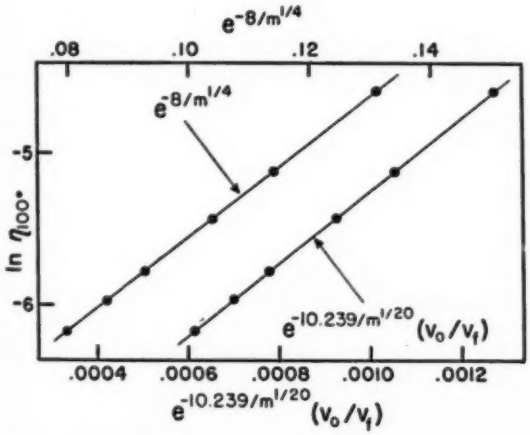


FIG. 1. *n*-Paraffin hydrocarbons at 100°C. Comparison of the molecular weight-free space equation with the logarithmic decrement rule for *n*-paraffins $m = 100$ through $m = 240$.

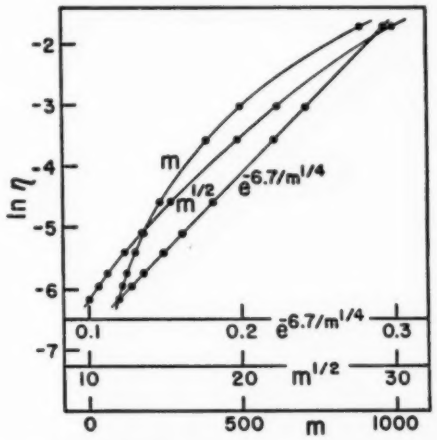


FIG. 2. *n*-Paraffin hydrocarbons at 100°C. Comparison of the linear and the square-root rules with the logarithmic decrement rule.

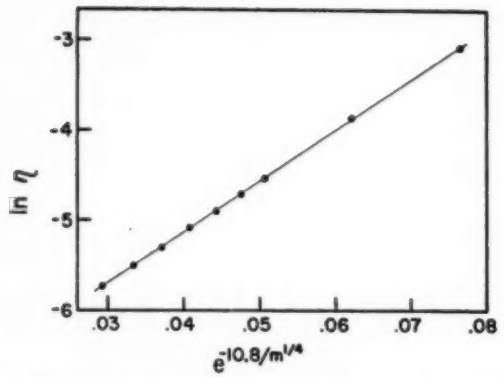


FIG. 3. *n*-Acetic esters at 50°C. Viscosities rectified according to logarithmic decrement rule.

8.0 to 6.7 in order to achieve linearity. This alters the position of the line, necessitating different values of α and γ. The equation then becomes:

$$\ln\eta_{100}=25.7\exp(-6.7/m^{1/4})-9.28, \qquad (5)$$

which reproduces the observed values from $m=100$ through $m=900$ with no value in error by more than 2.0 percent, see Fig. 2 and Table II. Comparison of Table II with Table I shows that Eq. (4) gives somewhat better agreement over its range, $m=100$ through $m=240$, than Eq. (5) does over the same range. This is due to the fact that in general, extending the range reduces the fidelity.

Figure 2, Table II, also presents a comparison of the linear and the square-root rules with the author's logarithmic decrement rule. It is apparent that over the range chosen, the linear or Dunstan rule fails to apply at all, whereas the square-root or Flory rule reproduces only the data from $m=100$ through $m=240$ with an accuracy equivalent to that of Eq. (5). It should be mentioned, however, that the square-root rule is better suited to treatment of data involving higher molecular weight substances than those illustrated by Fig. 2. This will be referred to again in a later section of the present paper.

Since the balance of this paper will deal with approximate rather than exact relationships, there is no

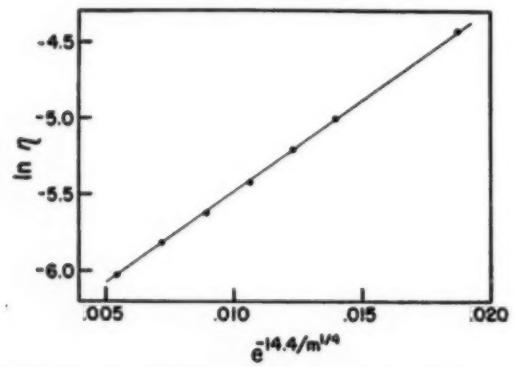


FIG. 4. *n*-2 Ketones at 50°C. Viscosities rectified according to logarithmic decrement rule.

necessity to be unduly concerned about the accuracy of the viscosity data under review. Thus, the *n*-alkane data used in Fig. 2 represent the entire range of our measurements, not just the "selected values" referred to in the previous papers. Likewise, viscosity measurements made on several other families of compounds which were taken with the Hoeppler precision viscometer rather than with the more accurate Ubbelohde viscometer will now be treated. The characterization of these compounds and their criteria of purity have been published elsewhere.[3] Other data taken from the literature are also included.

The viscosities of the members of any homologous series of compounds at fixed temperature may be rectified according to Eq. (2). This is best done by assuming a value for β and plotting $\ln\eta$ *vs* $\exp(-\beta/m^{1/4})$. If too low a value of β has been chosen, a curve that is concave upward will result. If too high a value has been chosen, the curve will be concave downward. The correct value of β can be quickly ascertained by making a

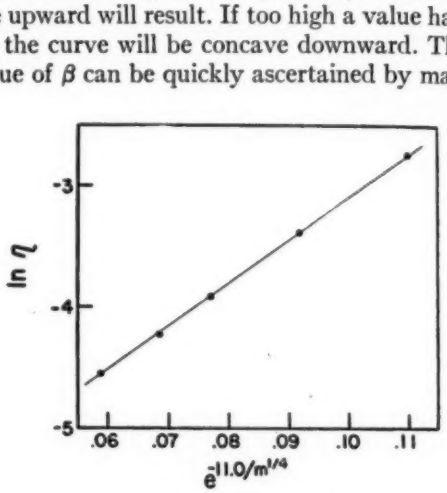


FIG. 5. Di-*n*-alkyl succinates at 100°C. Viscosities rectified according to logarithmic decrement rule.

few trial calculations. The results for six homologous series of compounds rectified in this manner are presented in Figs. 3–5 (data obtained in this laboratory), Fig. 6 (data of Flory[4]), and Figs. 7 and 8 (data of Fox and Flory[5]).

THE VISCOSITY OF MIXTURES OF LIQUIDS

The viscosity of mixtures of Newtonian liquids has been studied by numerous investigators. A brief reference to the principal contributions in this field may be found in a chapter by the author in Alexander's *Colloid Chemistry*.[6] It is apparent that most of the early work was devoted to a search for means of combining the viscosities or fluidities of the components in some manner that would express the viscosity or fluidity of the mixtures. Although this method of approach is useful in the case of two-component mixtures of normal

[3] A. K. Doolittle, Ind. Eng. Chem. **36**, 239 (1944).
[4] P. J. Flory, J. Am. Chem. Soc. **62**, 1057 (1940).
[5] T. G. Fox and P. J. Flory, J. Am. Chem. Soc. **70**, 2384 (1948).
[6] J. Alexander, *Colloid Chemistry* (Reinhold Publishing Corporation, New York, 1950), Vol. VII, pp. 161–162.

liquids having molecular weights of the same order of magnitude, it is not adequate to deal with multicomponent mixtures, and fails entirely when the difference between the molecular weights of the components becomes appreciable.

Perhaps the first paper to consider the law of mixtures in terms of the molecular weights of the components was that of Flory[4] who dealt with mixtures by his square-root rule, using the weight-average molecular weight of the components to represent the molecular weight of the mixtures. This procedure gave a very satisfactory fit of the data on decamethylene adipate polymers in molecular weight ranges 1400 to 14,000, but is not of general applicability either to mixtures containing lower molecular weight components or to homologous series of compounds of low molecular weight (Fig. 2). The author's approach to the treatment of the viscosity of mixtures of liquids and of polymer solutions may be regarded as an extension of the Flory treatment, although our work leading to this concept was done independently some years prior to Flory's.

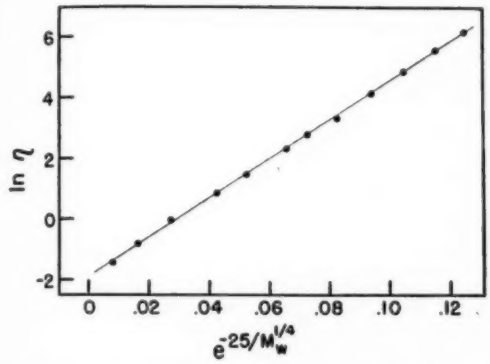


FIG. 6. Decamethylene adipate polymers at 109°C. Data of Flory rectified according to logarithmic decrement rule.

In the preceding section an approximation function was developed for rectifying the viscosities of homologous series of liquids expressed as functions of their molecular weights. We will now extend this treatment to apply to mixtures of liquids, and finally to solutions of resinous substances. This treatment merely involves a consideration of the methods of averaging molecular weights and the application of the appropriate weighting factor to each system studied.

THE WEIGHTED-AVERAGE MOLECULAR WEIGHT

The use of the terms "number-average," "weight-average," and "Z-average" has been common practice in high polymer work for a number of years, these terms having been first introduced by Kraemer and Lansing[7] in 1935. The number-average is the true or simple average, whereas all others are weighted averages. In the number-average, the contribution of the members of each size group to the over-all phenomenon is

[7] E. O. Kraemer, and W. J. Lansing, J. Phys. Chem. 39, 153 (1935); J. Am. Chem. Soc. 57, 1369 (1935).

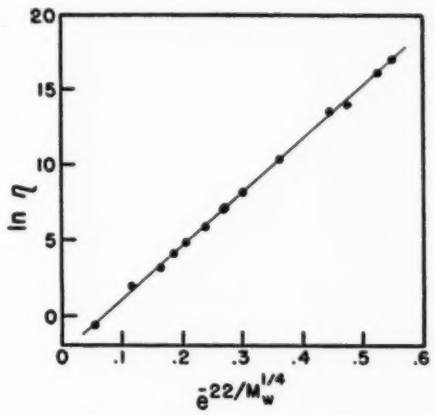


FIG. 7. Polyisobutylene fractions at 217°C. Data of Fox and Flory rectified according to logarithmic decrement rule.

uniform. In a weighted average, a different emphasis is placed on the proportionate contribution of the members of each size group to the total.

Consider, for example, an assortment of molecules of various sizes, in which m_1, m_2, m_3, etc. represent the mass of the molecules of each size group present, and n_1, n_2, n_3, etc., the number of molecules in each such group. The total number of molecules will then be Σn and the total mass, Σnm.

If equal emphasis is placed on the contribution of the individual molecules in each size group, as would be the case for example if we were counting the total number of terminal groups by means of titration with an appropriate reagent, the average molecular weight obviously would be the total mass divided by the total number of molecules. Thus we write the number-average,

$$\bar{M}_n=\Sigma\left[\frac{n_1}{\Sigma n}\cdot m_1+\frac{n_2}{\Sigma n}\cdot m_2+\frac{n_3}{\Sigma n}\cdot m_3+\right]=\frac{\Sigma nm}{\Sigma n}. \quad (6)$$

If, however, we measure some phenomenon in which equal emphasis cannot be placed on the contribution of

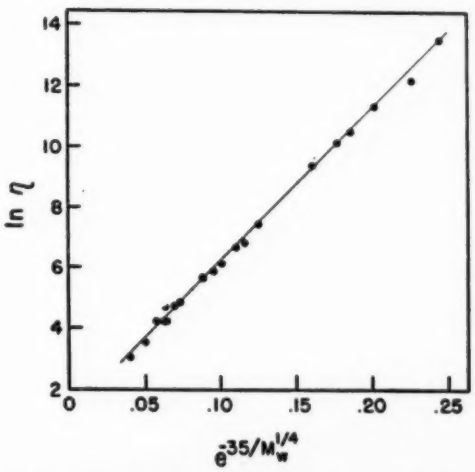


FIG. 8. Polystyrene fractions at 217°C. Data of Fox and Flory rectified according to logarithmic decrement rule.

TABLE III.

Chemical class	Temperature °C	Mol. wt. range	β
n-Paraffins	100	72–900	6.7
n-Acetic esters	50	88–312	10.8
n-2-ketones	50	58–170	14.4
Di-*n*-alkyl succinates	100	230–623	11.0
Decamethylene adipate polymers	109	738–20,850	25.0

the members of each size group, such as would be the case with diffusion or with intermolecular attraction, then it becomes necessary to assign a factor or "weighting" to the members of each size group to compensate for the unequal contribution of molecules of different sizes. Thus, if we weight the contribution of the members of each size group by a factor equal to the mass of a molecule of that group, we have the weight-average,

$$\bar{M}_w=\sum\left[\frac{n_1m_1}{\Sigma nm}\cdot m_1+\frac{n_2m_2}{\Sigma nm}\cdot m_2+\frac{n_3m_3}{\Sigma nm}\cdot m_3+\right]=\frac{\Sigma nm^2}{\Sigma nm}. \quad (7)$$

Another way of thinking of the same thing is to consider the hypothetical "average molecule" as a composite made up by adding together fractional parts of the molecular weights representing each size group present. The fraction by which the molecular weight of each size group is multiplied is determined by the particular method of weighting chosen. Thus, if equal emphasis is desired, the fraction is $n_i/\Sigma n$ or a number-fraction, and the average is called the number-average. If the weighting or emphasis is taken proportional to the molecular weight of each size group, the fraction becomes $n_im_i/\Sigma nm$. Since n_im_i is the weight of the ith size group, whereas Σnm is the total weight, the fraction is a weight-fraction, and this average is called a weight-average.

It is, of course, entirely possible to weight the contribution of each size group by some factor other than m^0 or m^1 as in the two averages cited above. In the case of the Z-average proposed by Kraemer and Lansing[7]

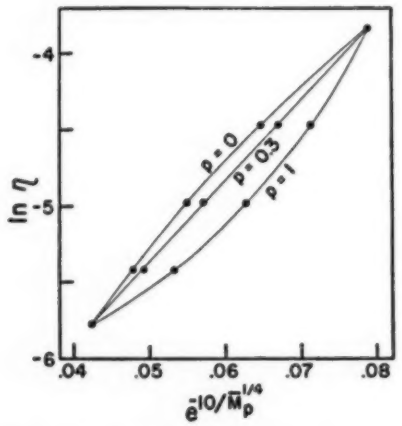


FIG. 9. *n*-Heptane and *n*-heptadecane at 50°C. Viscosities of mixtures rectified according to logarithmic decrement rule, showing influence of parameter, p.

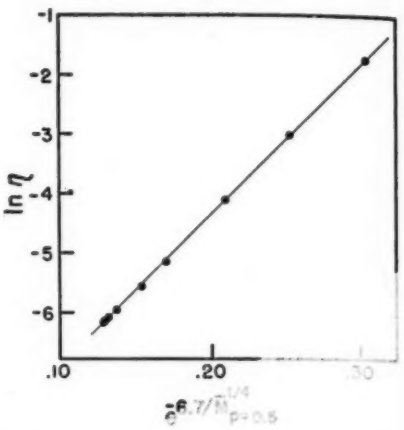


FIG. 10. *n*-Heptane and *n*-tetratriacontane at 100°C. Viscosities of mixtures rectified according to logarithmic decrement rule.

for use in evaluating sedimentation equilibrium data, the weighting factor used is m^2.

Not only may the function of m that truly represents the appropriate weighting factor be a complicated one, but its exact character may prove quite difficult or impossible to determine in certain instances. This is particularly so in the present study where we are dealing with approximation functions that do not have explicit physical significance. We propose, therefore, the use of a simple function, m^p, as the weighting factor, and will show that by assigning the appropriate value of the parameter, p, weighted-average molecular weights may be calculated that permit a wide variety of data on the viscosity of mixtures to be rectified by the same equations that apply to homologous series of pure liquids.

The general expression for the weighted-average molecular weight becomes, therefore:

$$\bar{M}_p=\sum\left[\frac{n_1m_1^p}{\Sigma nm^p}\cdot m_1+\frac{n_2m_2^p}{\Sigma nm^p}\cdot m_2+\frac{n_3m_3^p}{\Sigma nm^p}\cdot m_3+\right]=\frac{\Sigma nm^{p+1}}{\Sigma nm^p}. \quad (8)$$

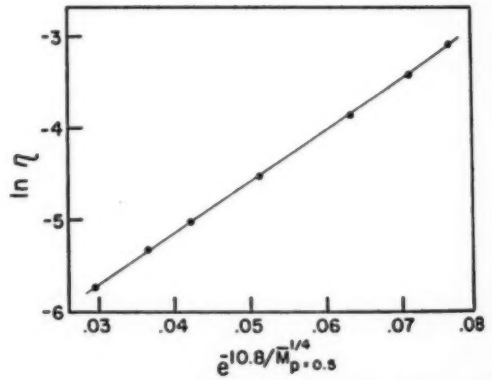


FIG. 11. Ethyl acetate and *n*-octadecyl acetate at 50°C. Viscosities of mixtures rectified according to logarithmic decrement rule.

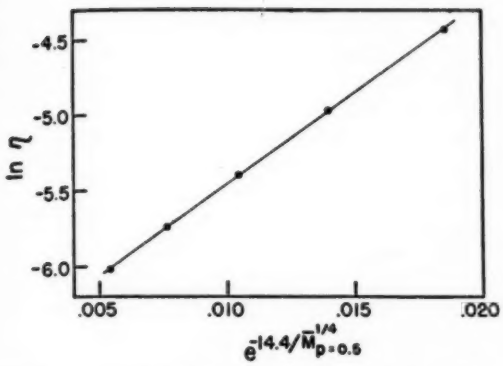


FIG. 12. Acetone and methyl-*n*-nonyl ketone at 50°C. Viscosities of mixtures rectified according to logarithmic decrement rule.

From this expression, it is clear that the number-average is defined when $p=0$, the weight-average when $p=1$, and the Z-average when $p=2$. Intermediate weighted averages are defined by fractional values of the parameter, p,

If we have a mixture of only two components and their molecular weights are represented by m and M, respectively, and if w_2 represents the weight-fraction of component M in the mixture, we may simplify Eq. (8) to:

$$\bar{M}_p=m+(M-m)/[1+(M/m)^{1-p}(1-w_2)/w_2]. \quad (9)$$

ADAPTATION OF THE LOGARITHMIC DECREMENT EQUATION TO THE VISCOSITY OF MIXTURES

The approximate molecular weight function,

$$\ln\eta=\alpha\exp(-\beta/\bar{M}_p^{1/4})+\gamma \quad (10)$$

taken from Eq. (2), may be applied to data on the viscosity of mixtures exhibiting Newtonian flow by using average molecular weights calculated according to Eqs. (8) or (9), for which an appropriate value of the parameter, p, has been chosen.

Before the parameter, p, can be determined, however, it is first necessary to establish the value of β in Eq. (10). If the components of the mixture are members of an homologous series of liquids, and viscosity data are at hand on a number of the pure members of the series at the desired temperature, the value of β may be determined by trial as explained in a prior section. If such data are not available, choice of an appropriate value may be made by comparison with values for homologous series of compounds of similar chemical types. In general, a precise value for β is not essential. Thus the values of β in Table III are established from the data so far presented.

Once the value of β has been chosen, the value of the parameter, p, is obtained by trial calculations. That is, a value is assigned and the average molecular weights calculated for the various mixtures under study. The logarithms of the viscosities are then plotted against $\exp(-\beta/\bar{M}_p^{1/4})$. Too low a value of p will give a curve concave downward, whereas too high a value of p will make the curve concave upward. For the appropriate value of p, a straight line will result. Figures 9–15 listed in Table IV illustrate the conformance attained in the case of a number of systems that we have studied.

The value of the parameter, p, is seen from the tabulation in Table IV to vary with the nature of the sub-

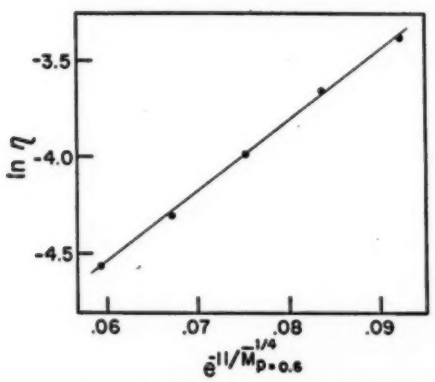


FIG. 13. Di-*n*-butyl and di-*n*-dodecyl succinate at 100°C. Viscosities of mixtures rectified according to logarithmic decrement rule.

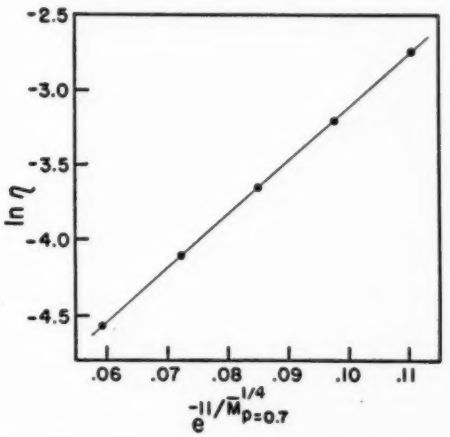


FIG. 14. Di-*n*-butyl and di-*n*-octadecyl succinate at 100°C. Viscosities of mixtures rectified according to logarithmic decrement rule.

TABLE IV.

Figure	Components	Temperature °C	Parameter	β
9	*n*-Heptane and *n*-heptadecane (Note influence of $p=0$ and $p=1$ for this system)	50	0.3	10.0
10	*n*-Heptane and *n*-tetrahexacontane	100	0.5	6.7
11	Ethyl acetate and *n*-octadecyl acetate	50	0.5	10.8
12	Acetone and methyl *n*-nonylketone	50	0.5	14.4
13	Di-*n*-butyl and di-*n*-dodecyl succinates	100	0.6	11.0
14	Di-*n*-butyl and di-*n*-octadecyl succinates (Data for the above obtained in this laboratory)	100	0.7	11.0
15	Decamethylene adipate polymers $m=1402$, $M=14{,}110$ (Data of Flory (see reference 4))	109	1.0	25.0

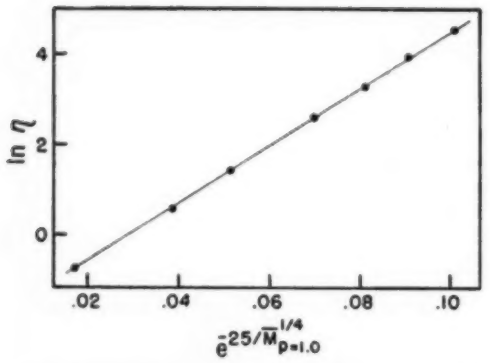


FIG. 15. Decamethylene adipate polymers, $m=1402$ and $M=14{,}100$, at 109°C. Data of Flory rectified according to logarithmic decrement rule.

stances, the spread in molecular weights involved, and the order of magnitude of the molecular weights of the components. Mixtures of high molecular weight substances with a considerable spread between the molecular weights of the two components generally seem to require a value of $p=1$, which accounts for the successful use of weight-average molecular weights by Flory and others dealing with mixtures of polymers. On the other hand, the closer the molecular weights of the components are to each other, the nearer p approaches zero, since the influence of the higher molecular weight component is diminished and the average approaches the simple or number-average molecular weight.

ADAPTATION OF THE LOGARITHMIC DECREMENT EQUATION TO THE VISCOSITY OF SOLUTIONS OF HIGH POLYMERS

The Viscosity of Solutions of Resinous Substances

The viscosity of solutions of macromolecular materials in appropriate solvents is a subject of immense technological importance in connection with both the manufacture and the use of resins in industry. The study that has culminated in the present series of four papers on the viscosity of Newtonian liquids was initiated about fifteen years ago in order to afford a better understanding of the viscosity phase relationships of resin solutions, first described in connection with nitrocellulose by the author[8] in 1938 and in connection with "Vinylite" resins by Quarles[9] in 1943. For the most part, the concentrated solutions and suspensions of resinous substances that are useful in industry in adhesives, surface coatings, and synthetic fibers do not exhibit Newtonian flow, and their viscosity characteristics are largely explained by their departure from Newtonian flow rather than by the fundamental principles of Newtonian flow itself. Thus, the more complicated examples of flow result from the simultaneous effect of several phenomena that are superimposed on one another. Examples of superimposed effects are elastic resistance arising from gelation, higher order cohesive forces such as those caused by local crystallization, or merely the mechanical resistance of particles resulting from close packing. It was early recognized by us, therefore, that little progress could be expected in elucidating the anomalies of flow exhibited by many concentrated resin solutions until a better understanding was gained of the simple fundamental effect upon which the anomalous behavior was superimposed. This led us eventually to the study of the normal paraffin hydrocarbons and of several homologous series of other Newtonian liquids, the results of which have been set forth in this series of papers.

The investigation of the flow behavior of solutions of macromolecular substances by others has followed two quite different lines of approach. On the one hand, the flow of dilute solutions has been treated hydrodynamically as a special case of the theory of suspensions, and a considerable contribution has been made by the theoretical men who have followed this line of investigation. On the other hand, numerous investigators have attempted to set up empirical relations between viscosity and concentration of solute for a large number of systems. The latter approach, although useful in individual cases, does not contribute much to our understanding of the flow behavior of solutions because it does not differentiate between the effects of the different factors that combine to produce the over-all effect observed. For a résumé of the principal contributions to this aspect of the subject, the reader is again referred to the chapter by the author in Alexander's *Colloid Chemistry*.[10]

The treatment of solutions of high polymers that we propose involves a slight extension and some simplification of the ideas advanced in the prior section on mixtures. The extension arises from the necessity of considering the solvent as a member of the same homologous series as the resinous component or solute. In most cases this offers no complications, but in the case of solvent and solute of quite dissimilar types, the appropriate weighted-average molecular weights may differ from those that apply to members of a true homologous series. The simplification arises from the fact that, in solutions of high polymers, the spread in molecular weight between solvent and solute is so great that such solutions may be treated as special cases of the more general rules that apply to mixtures. Thus, if solvent and solute do belong to the same homologous series, the great spread between their molecular weights usually calls for a value of *one* for the parameter, p, of Eq. (9). Under such circumstances Eq. (9) reduces at once to:

$$\bar{M}_{p=1}=\bar{M}_w=m+(M_w-m)\cdot w_2. \qquad (11)$$

If, however, solvent and solute are of quite different chemical families, a value of p somewhat less than *one*

[8] A. K. Doolittle, Ind. Eng. Chem. **30**, 199 (1938).
[9] R. W. Quarles, Ind. Eng. Chem. **35**, 1033 (1943).

[10] See reference 6, pp. 156–157.

often has been found to apply. If this value is 0.5 or less, a different method of simplification of Eq. (9) may be employed. Thus, we may write:

$$\bar{M}_p = m + \frac{M_p - m}{1 + (M_p/m)^{1-p}(1-w_2)/w_2}$$

$$= \frac{M_p - m}{(M_p/m)^{1-p}} \cdot \frac{w_2}{1 - w_2(1 - 1/(M_p/m)^{1-p})}. \quad (12)$$

If M/m is a sufficiently large number, $1/(M/m)^{1-p}$ becomes negligible in comparison with unity for values of p equal to or less than about 0.5. This simplification is valid for low values of w_2, but differs appreciably from the true function as w_2 approaches unity. The simplified expression is:

$$\bar{M}_{p \leq 0.5} = m + \frac{M_p - m}{(M_p/m)^{1-p}} \cdot \frac{w_2}{1 - w_2}. \quad (13)$$

Dependence on Molecular Weight

If the molecular weight of the polymer is known, Eqs. (11) or (13) (whichever is appropriate to the system studied) may be applied as written. Figure 16 presents Flory's data on solutions of decamethylene adipate polymer ($M_w = 16{,}070$) in diethyl succinate ($m = 174.2$) at 79°C, rectified according to the logarithmic decrement rule using the simplification indicated in Eq. (11). Attention is particularly directed to the good conformance of the data at very low concentrations of solute. The same data plotted according to the square-root rule, Fig. 17, fail to conform in this region.

Dependence on Concentration

In most industrial applications where viscosity control or blending operations are carried out, the molecular weight is not known. In such cases it is desirable to rectify viscosity as a function of solute concentration without reference to molecular weight. Equation (10)

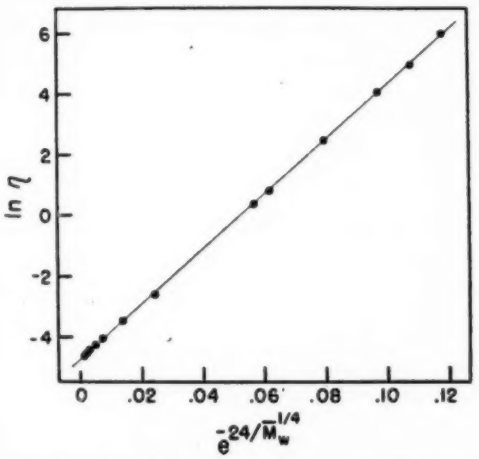


FIG. 16. Decamethylene adipate polymer in diethyl succinate at 79°C. Data of Flory rectified according to logarithmic decrement rule.

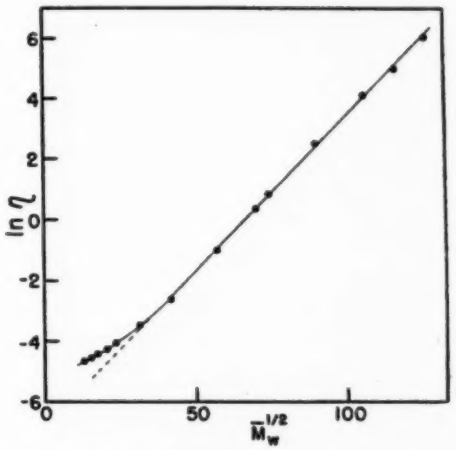


FIG. 17. Decamethylene adipate polymer in diethyl succinate at 79°C. Data of Flory rectified according to square-root rule.

is readily modified for this purpose by further simplification of Eqs. (11) and (13). Before proceeding with this step, however, it is necessary to consider certain ways of expressing concentration in solutions. In industrial work, concentration is almost always expressed as percent solids or grams solute per 100 g of solution. In scientific work, however, it is often useful to express concentration in grams solute per 100 g solvent. The symbol c is used in either case, with the appropriate units indicated. Now, for any given system of solvent and solute, both the $(M_w - m)$ of Eq. (11) and the $(M_p - m)/(M_p/m)^{1-p}$ of Eq. (13) are fixed. Likewise, $w_2 = c/100$ g solute per gram of solution, and $w_2(1-w_2) = c/100$ g solute per gram of solvent. Consequently, both Eqs. (11) and (13) may be written:

$$\bar{M}_p = m + kc \quad (14)$$

for the particular values $p = 1$ and $0 < p \leqq 0.5$, provided the unit of concentration is properly defined.

For the special cases just discussed, Eq. (10) could be written:

$$\ln\eta = \alpha \exp[-\beta/(m+kc)^{1/4}] + \gamma, \quad (15)$$

but a further simplification is desirable and may be made by translating the axes defining $\ln\eta = f(\bar{M}_p)$ a distance of m units to the right. The same curve is then defined with respect to the new position of the axes by $\ln\eta = f(kc)$, since $\bar{M}_p = m + kc$. Therefore, we may write for high polymer solutions,

$$\ln\eta = \alpha \exp[-\beta/(kc)^{1/4}] + \gamma = \alpha \exp(-K/c^{1/4}) + \gamma, \quad (16)$$

for the special cases where either the weight-average molecular weight applies exactly or an average approaching the number-average molecular weight applies approximately. In the former case c is in grams solute per hundred grams solution, α is the same as in Eq. (10), and γ is $\ln\eta_{solv}$. In the latter case, however, an approximation is made that introduces an appreciable deviation toward the higher concentration end. In order to rectify the viscosity as a function of concen-

tration in this case, a slight adjustment to take care of this deviation must be made in the β of Eq. (10), requiring somewhat different values of α and γ in Eq. (16).

Equation (16) is of wide applicability to polymer solutions that exhibit Newtonian flow, and has been known and used by us for many years. In fact, the logarithmic decrement rule, first described in this paper, was originally discovered some fifteen years ago in the form of Eq. (16) by studies based on the viscosities of solutions of high polymers.

Figure 18 illustrates an application of Eq. (16) with data on the viscosities of acetone solutions of nitrocellulose at 20°C. In this case the value of the parameter, p, happens to be 0.5, so that concentration is expressed as grams solute per 100 g solvent. The agreement is good from 1 to 20 percent solids. (1 to 25 g/100 g solvent).

Better agreement over a broader range of concentrations is obtained in the case of the acetone solutions of nitrocellulose by use of Eqs. (9) and (10) which, of course, presupposes a knowledge of the molecular weight of the solute. Figure 19 illustrates this case, where the value of $M_{0.5}=23{,}250$ is obtained by multiplying the number average, $M_N=15{,}500$, by three-halves. (The value of the number-average molecular weight was determined by the osmotic pressure method.) When the viscosities of the acetone solutions of nitrocellulose at 20°C were rectified according to Eqs. (9) and (10), all points from zero to 35 percent solids fell on the line as compared to only the values from 1 to 20 percent solids in the previous case. For comparison, the Dunstan (linear) and Flory (square-root) rules are plotted on Fig. 19, using the same values of $\bar{M}_p$ throughout.

This paper completes the series of four articles published under the title "Studies in Newtonian Flow," inasmuch as the dependence of the viscosity of liquids on temperature, on free-space, on molecular weight in homologous series or mixtures of homologs, and on concentration in polymer solutions has been discussed. At a future date it is planned to offer a paper for publication in which the theoretical implications of these observations will be discussed.

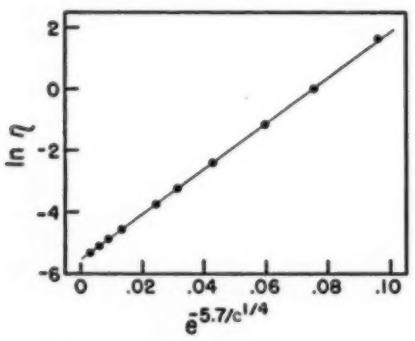


FIG. 18. Dependence on concentration. Half-second nitrocellulose in acetone at 20°C. Viscosities of high polymer solutions rectified according to logarithmic decrement rule (concentration in grams solute per 100 g solvent).

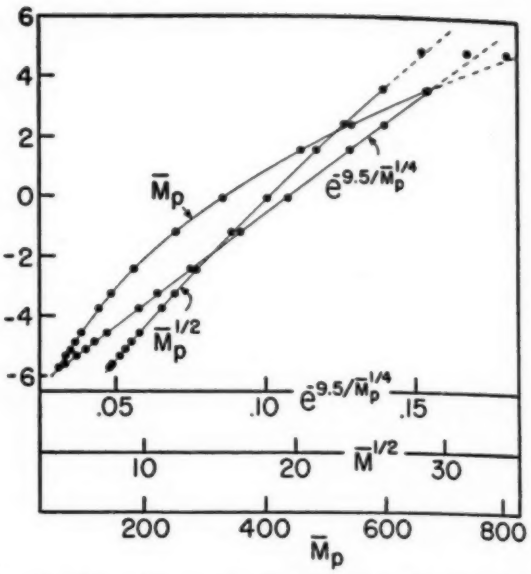


FIG. 19. Dependence on molecular weight. Half-second nitrocellulose in acetone at 20°C. Viscosities of high polymer solutions rectified according to (1) linear rule, (2) square-root rule, (3) logarithmic decrement rule.

SUMMARY

1. The free-space—molecular weight equation developed in a previous paper is here modified by substituting for the molecular weight—free-space variable its approximate equivalent in terms of molecular weight at constant temperature. Further minor adjustment of the constants gives the following generalized approximate molecular weight equation applying to homologous series of pure liquids at constant temperature:

$$\ln\eta=\alpha\exp(-\beta/m^{1/4})+\gamma.$$

2. The above expression is shown to be applicable to mixtures of liquids and to solutions of polymers by substituting for the molecular weight, m, an appropriate weighted-average molecular weight, $\bar{M}_p$, representing all components of the mixture or solution.

3. For the special cases where the weighting factor parameter, p, is either equal to *one*, or is equal to or less than 0.5, the mixture equation may be further modified to represent the relation between viscosity and concentration in polymer solutions. The equation in this form becomes:

$$\ln\eta=\alpha\exp(-K/c^{1/4})+\gamma.$$

ACKNOWLEDGMENT

The author desires to express his gratitude to the Carbide and Carbon Chemicals Company for supporting this project to completion, and to the vice-president in charge of research, G. A. Perkins, for encouragement and advice throughout the course of the work. Many members of the Research Staff contributed to this study, but those most immediately concerned were R. H. Peterson and R. W. Callard. The author's wife, Dortha Bailey Doolittle, assisted with the literature search and with the translation of many articles from foreign languages.

Metallurgical Applications of X-Ray Fluorescent Analysis

P. K. Koh and Betty Caugherty
Research Laboratories, Allegheny Ludlum Steel Corporation, Brackenridge, Pennsylvania
(Received October 12, 1951)

Techniques are described for fluorescent analysis using an x-ray Geiger counter goniometer. Consideration is given to the influence of potential variations, specimen preparation, and thickness of the diffracting crystal.

The observed intensity of fluorescent radiation from a number of common commercially pure metallic elements was compared under a constant primary radiation that can be closely maintained in practice.

Intensities of fluorescent spectra were measured on constituents of Fe-Ni, Fe-Cr, and Fe-Ni-Cr alloys. Calibration curves correlating line intensity with weight percentage are included for binary Fe-Ni and Fe-Cr alloys.

The thickness beyond which the fluorescent intensity becomes constant has been determined for pure iron, nickel, and chromium. Up to the above thickness for each element an empirical plot of fluorescent intensity against thickness of layer has been established. These empirical plots can be applied to measure the thickness of very thin coatings.

Illustrations are given indicating the applications of the fluorescent analytical technique to the study of oxidation and diffusion processes in metallic systems.

An aluminum holder is used for analyzing small quantities of powder derived from chemical or electrolytic extraction of certain phases from alloys. Fluorescent analysis on the extracted phase supplements the diffraction data when a wide solid solubility exists.

INTRODUCTION

THE principle of chemical analysis by secondary x-ray radiation has been discussed fully by Von Hevesy,[1] Glocker,[2] Schreiber,[3] Freidman,[4] and others. The characteristic secondary radiations of a metallic or nonmetallic specimen having wavelengths proportional to the inverse square of atomic number of the containing elements, are excited by the primary radiation from a separate x-ray tube. The secondary radiations are diffracted through a certain known set of atomic planes of a crystal and then energize a Geiger counter mounted on a goniometer. By such means the wavelengths as well as intensities of the resultant secondary radiations can be readily analyzed.

Earlier workers used both primary and secondary x-ray radiations under vacuum for their analysis. Since that time advances have been made in elevating the power output of sealed-off x-ray tubes, and in improving the performance of Geiger counters. As a result commercial units are available in this country with both primary and secondary x-ray beams in air. This change of design simplifies the constructional and operational problems but also limits the applicability of this analytical method to the chemical elements having atomic numbers not less than 22 (titanium).

The excitation efficiency of chemical elements is a function of the atomic number. The resultant secondary characteristic radiations are of various wavelengths. Absorption coefficients in all media are dependent upon wavelength. At present classical theory does not permit a theoretical calculation of the mass absorption coefficient of heterogeneous primary radiation. In alloys there are added complications due to the absorption edge effect and the presence of foreign impurities. Thus the intensity of fluorescent radiation of a substance could be considered as a function of numerous variables such as geometrical set-up, excitation efficiency of the element, air absorption, absorption by the specimen of both primary and secondary radiations, mutual excitation, effect of foreign impurities, absorption by the bent mica crystal, absorption of the Geiger counter window, sensitivity of the Geiger counter to radiations of various wavelengths, etc. Compton[5] has derived the equation for fluorescence yield, and Glocker[2] has made the line intensity calculation in a few selective cases for determining the proportion of an element in a mixture. However, a quantitative analysis of the fluorescence yield by secondary radiation for elements in practical metallurgical alloys remains to be solved.

A survey of possible applications and limitations by this secondary radiation analysis to relevant metallurgical problems is attempted in the following paragraphs with brief experimental data.

TECHNIQUE

In the present investigation a General Electric SPG Fluorescent Analyzer was used in conjunction with the General Electric XRD-3 Spectrometer Goniometer. The primary radiation from a tungsten-target Machlett AEG-50T tube irradiates a ¾-inch × ¾-inch sample area. The fluorescent radiations are diffracted through a bent mica crystal and detected by a Geiger counter spectrometer. Readings were obtained by scaling at 1 and ½ percent probable error (4096 and 16,384 total counts,

[1] G. Von Hevesy, *Chemical Analysis by X-Ray and Its Application* (McGraw-Hill Book Company, Inc., New York, 1932).
[2] R. Glocker and H. Schreiber, Ann. Physik **85**, 1089–1102 (1928).
[3] H. Schreiber, Z. Physik **58**, 619–650 (1929).
[4] H. Friedman and L. S. Birks, Rev. Sci. Instr. **19**, 323–330 (1948).
[5] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Company, Inc., New York, 1935).

TABLE I. Effect of surface preparation on line intensity.

Specimen	Composition	Description	% Variation of K-alpha intensity Cr	Fe	Remarks
A	17 Cr-Fe	50 grit sand belt finish } cold rolled and annealed	0	0	ASTM grain size, 7 to 8
A	17 Cr-Fe	00 emery paper finish } cold rolled and annealed	+4.55	−1.83	
A	17 Cr-Fe	00 finish+2 min electrolytic etch in NaCN } cold rolled and annealed	+6.22	−0.97	
B	17 Cr-Fe	Cold rolled and annealed+1 hour at 2300°F, air cooled, 00 emery paper finish	−2.38	+0.66	ASTM grain size, 1 or larger

respectively) at 50 kv and 35 ma. Each reading represents the average of several runs.

An exploratory study was made by varying details in specimen preparation and in operating conditions. The following paragraphs describe the experimental results obtained.

Under a constant primary radiation and the same experimental set-up Table I shows that the smoothness of surface finish, grain size, or residual stresses of a metallic alloy do not greatly affect the line intensity of a characteristic radiation.

Table II shows the variation on changing tube potential of the fluorescent line intensity of an element in a binary alloy with respect to that of the unalloyed metal and also of the ratio between the line intensity of two elements in the binary alloy. It is evident from the tabulation that a comparison of line intensity measurement of elements in alloys should be made under the same operating conditions.

An inadequate warm-up period for the electronic tubes in the counting circuit usually causes a slight decrease in the line intensity.

TABLE II. Effect of tube potential on line intensities in binary alloys.

Kv	$Fe_{K\alpha}$ alloy/$Fe_{K\alpha}$	$Ni_{K\alpha}$ alloy/$Ni_{K\alpha}$	$Fe_{K\alpha}$ alloy/$Ni_{K\alpha}$ alloy
Specimens: 50 Ni-50 Fe, electrolytic iron and carbonyl nickel filament current =34.5 ma.			
20	0.655	0.265	2.66
25	0.650	0.283	2.12
30	0.672	0.303	1.87
35	0.666	0.317	1.67
40	0.672	0.333	1.57
45	0.677	0.341	1.48
50	0.671	0.345	1.44
Specimens: 79 Ni-21 Fe, electrolytic iron and carbonyl nickel filament current =34.5 ma.			
25	0.345	0.614	0.509
30	0.347	0.636	0.464
35	0.359	0.660	0.438
40	0.366	0.651	0.428
45	0.379	0.679	0.419
50	0.382	0.673	0.408
Specimens: 49 Cr-51 Fe, electrolytic chromium, electrolytic iron filament current =34.5 ma.			
Kv	$Fe_{K\alpha}$ alloy/$Fe_{K\alpha}$	$Cr_{K\alpha}$ alloy/$Cr_{K\alpha}$	$Fe_{K\alpha}$ alloy/$Cr_{K\alpha}$ alloy
20	0.335	0.528	1.47
25	0.342	0.542	1.45
30	0.361	0.580	1.45
35	0.366	0.576	1.46
40	0.373	0.562	1.48
45	0.382	0.578	1.48
50	0.383	0.583	1.47

The purity of the mica crystal affects the interplanar space of the diffracting atomic planes, which in turn affect the two-theta value of the peak fluorescent line intensity as it is recorded by the Geiger counter goniometer. The orientation of the diffracting atomic plane with respect to the cleavage plane of the mica crystal also differs individually. A slightly different setting is necessary for the maximum diffraction efficiency when a diffracting mica crystal is changed.

Table III shows the experimental data obtained on mica crystals of three different thicknesses on elements whose characteristic radiations cover the useful two-theta range of goniometer. In general it can be stated that for fluorescent radiations of long wavelength the thinner mica crystal diffracts more efficiently than the thicker one, while for shorter wavelengths, the thicker mica crystal does better.

RELATIVE INTENSITY OF FLUORESCENT CHARACTERISTIC RADIATIONS FROM SOME COMMERCIALLY PURE METALLIC ELEMENTS

The readings in Table IV were obtained under the same operating conditions on commercially pure metallic elements available in our laboratory. As shown in Table III, the relative order of intensity from various elements will change greatly if a diffracting mica crystal of appropriate thickness is used. In general, it can be stated that the *K*- or *L*-radiations become weaker at the extreme ends of lower or higher atomic numbers. While the fluorescent analysis is incapable of determining the percentage of aluminum in the aluminum alloys and hardly accurate for the percentage of titanium in the titanium alloys, the method is quite sensitive for the detection of small amounts of alloying elements, such as iron and manganese. In the same sense the fluorescent analysis is useful for the identification of impurities or segregates in commercial alloys.

INTENSITY MEASUREMENT OF THE FLUORESCENT RADIATION FROM AN ELEMENT IN BINARY AND TERNARY ALLOYS

In the absence of a workable line intensity calculation for quantitative determination of an element in an alloy, Friedman and Birks[4] made empirical plots of

TABLE III. Effect of the thickness of mica crystal on line intensities.

Element	Atomic number	Characteristic radiation[a] Line	2θ	Relative line intensity Thin mica 0.0007 in. thick	Medium mica 0.0014 in. thick	Thick mica 0.0042 in. thick
Titanium	22	K-alpha	133.2°	1	0.263	0.038
Chromium	24	K-alpha	99.7°	1	0.549	0.041
Iron	26	K-alpha	80.6°	1	0.742	0.240
Copper	29	K-alpha	61.9°	1	0.881	0.639
Molybdenum	42	K-alpha	27.4°	1	1.600	2.925
Cadmium	48	K-alpha	20.6°	1	1.813	4.613
Bismuth	83	L-alpha one	44.8°	1	1.073	1.527

[a] Tube potential 50 kv, filament current 35 ma.

measured intensity of FeKα radiation of an alloy relative to that of 100 percent iron *vs* the atomic percentage of iron in the Fe-Al, Fe-Ni, and Fe-Ag alloys. In the present investigation the same empirical plots were extended to Fe-Cr and Fe-Ni alloys but on the more convenient weight percentage basis (Fig. 1). The plots deviate largely from a straight line relationship because of the phenomena of mutual excitation and absorption. For general and routine analysis it is convenient to resort to calibration curves of fluorescent intensity *vs* weight percentage of an element in binary, ternary, or commercial complex alloys of compositions similar to the actual sample.

Table V shows the fluorescent line intensity *vs* weight percentage of alloying elements in Fe-Cr-Ni ternary alloys.

MINIMUM EFFECTIVE THICKNESS FOR FLUORESCENCE AND MEASUREMENT OF THICKNESS OF A METALLIC ELEMENT

The minimum effective thickness of a metallic element for fluorescence is considered to be the thickness which produces the maximum fluorescent yield by that element under the specified operating conditions of the primary tube. With a constant primary intensity any thickness under the minimum effective thickness produces a lesser intensity of the characteristic radiation, but any greater thickness produces no greater intensity.

TABLE IV. Intensity of fluorescent characteristic radiations from some commercially pure metallic elements relative to iron K_α.

Element	Atomic number	Atomic weight	Relative intensity of fluorescent lines $K_{\alpha 1}+K_{\alpha 2}$	K_β	$L_{\alpha 1}$	$L_{\beta 1}$	$L_{\beta 2}$
Ti	22	47.90	0.071	0.031			
Cr	24	52.01	0.317	0.069			
Mn	25	54.93	0.558	0.150			
Fe	26	55.85	1.000	0.263			
Co	27	58.94	1.362	0.321			
Ni	28	58.69	1.672	0.423			
Cu	29	63.54	1.564	0.346			
Zn	30	65.38	1.747	0.363			
Mo	42	95.95	0.634	0.091			
Cd	48	112.41	0.072	0.011			
Sn	50	118.70	0.037	0.008			
Ta	73	180.88			0.390	0.242	0.115
W (Powder)	74	183.92			0.208	0.165	0.059
Pt	78	195.23			0.270	0.153	0.079
Pb	82	207.21			0.230	0.168	$(L_{\beta 1}+L_{\beta 2})$
Bi	83	209.00			0.260	0.190	$(L_{\beta 1}+L_{\beta 2})$
U[a]	92	238.07			0.067	0.014	0.013

[a] Undersized sample.

TABLE V. Line intensity *vs* weight percentage of alloying elements in Fe-Cr-Ni ternary alloy.

Alloy	Cr$_{K\alpha}$ alloy/ Cr$_{K\alpha}$	Ni$_{K\alpha}$ alloy/ Ni$_{K\alpha}$	Fe$_{K\alpha}$ alloy/ Fe$_{K\alpha}$
19.2 Cr, 8.9 Ni, 71.9 Fe	0.317	0.055	0.686
17.5 Cr, 10.4 Ni, 72.1 Fe	0.296	0.067	0.644

No exact data on the minimum effective thickness for fluorescence of common metallic elements are available in the literature. Glocker[2] estimated the thickness to be in the order of $\frac{1}{10}$ mm.

For practical applications the minimum effective thickness of a substance means the maximum penetration by the primary radiation within which a change in percentage of an element can be detected by its change in intensity of fluorescent characteristic radiation. But if the same chemical change should occur further inward from the surface than its minimum effective thickness, the change cannot be evidenced. It is evident that a clear knowledge of the minimum effective thickness of a metallic element is of considerable importance in studies of diffusion, oxidation, surface adsorption, and plating.

For example, the minimum effective thicknesses of Cr, Ni, and Fe are determined in this investigation (Fig. 2) by building up successive layers of each metal on a piece of aluminum by electrolytic plating. In each case a low temperature, a low current density, and a suitable *p*H range of the electrolyte are used to insure a tight and uniform coating. The thickness is determined by weighing the deposit on a known surface area. For very thin coatings such as hard chromium plating on tool steels or metallized surfaces on plastics the thickness of the coating can be determined from empirical curves such as Fig. 2.

OXIDATION FILM STUDY BY FLUORESCENT ANALYSIS

As a result of the very low value of the mean effective thickness of a metal or an alloy as discussed in the preceding section it follows naturally that the x-ray

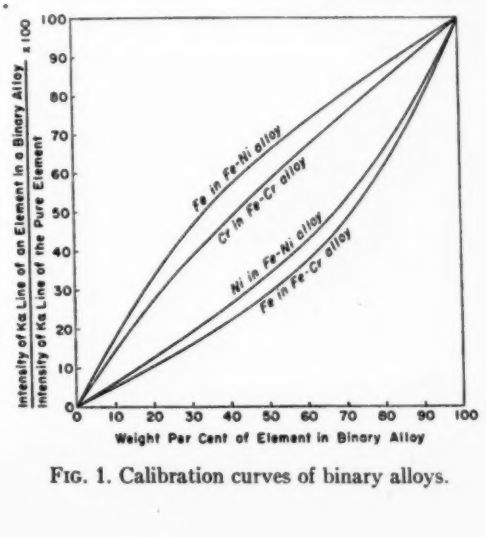


FIG. 1. Calibration curves of binary alloys.

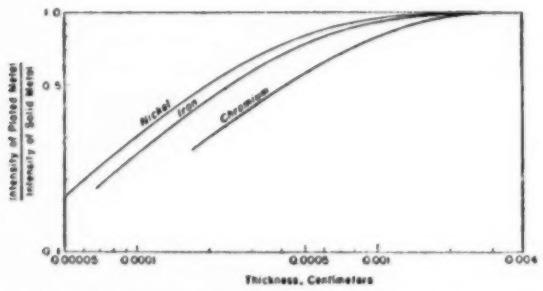


FIG. 2. Effect of sample thickness on the intensity of secondary radiation. Samples were plated on aluminum.

fluorescent analysis serves as an effective means for observing chemical changes on or very near the surface. Oxide films formed on alloys at elevated temperatures serve as an interesting example.

In the past, electron diffraction[6] was used because of the shallow penetration of the soft electron beam. In the following examples on oxidation of Fe-Ni, Fe-Cr, and Fe-Cr-Ni alloys, it is clearly shown that the same mechanism of oxidation and diffusion and the same order of formation of the several layers of films can be observed by the x-ray fluorescent analysis.

For oxidation study by x-ray fluorescence specimens of four different alloys were heated in air at 2000°F for total accumulated time periods of 1 minute, 6 minutes, 30 minutes, 1 hour, and 3 hours. During air cooling from the furnace temperature, a superficial layer of film peeled off the metallic specimen. At the same time a new adherent oxidized film formed and replaced the film that peeled off. Both the peeled film and the oxidized surface in every case were subjected to fluorescent analysis, and if necessary, also to x-ray diffraction. The results of fluorescent analysis, in this case, represent the gross average composition of a layer of thickness less than 0.003 cm, the layer nearest to the surface contributing most strongly to the results. These results are by no means as selective as those obtained by electron diffraction, but they give information on chemical changes on the surface layer, which are not shown by electron diffraction.

An intensity count was made on an unoxidized specimen along with each of the oxidized specimens as a check on the stability of operating conditions.

When a 17 Cr-Fe alloy is oxidized in air at 2000°F, the initial oxidized surface which is adherent to the underlying metallic surface becomes at first richer in chromium content than that in the original alloy. On prolonged oxidation a superficial layer of scale peels off, and a renewed adherent oxidized layer is exposed which subsequently falls off. As shown in Table VI, the first peeled scale has chromium content approximately the same as its original alloy content. The successive peeled scale, however, becomes greatly impoverished in chromium. The same trend is followed in the adherent oxidized surface which, although initially enriched in chromium, transforms at a later stage into mostly iron oxide. The Debye diffraction pattern of peeled scale after 30 minutes at 2000°F in air shows strong lines of Fe_3O_4, less intense lines of Fe_2O_3 and weak lines of Cr_2O_3.

On oxidizing a 50 Ni-50 Fe alloy at 2000°F in air a blistered scale appears immediately on air cooling after the briefest moment in the furnace. The oxidized layer next to the metallic surface renewed itself repeatedly and rapidly. Table VI shows that the adherent oxidized layer passes through a discontinuous transformation in chemical composition after 3 hours at 2000°F in air. On oxidizing the alloy specimen for time periods less than 3 hours the adherent layer becomes steadily impoverished in nickel content in comparison with its original alloy content. But after oxidation for 3 hours the adherent oxidized layer suddenly becomes enriched in nickel content. The peeled scales after various time periods of oxidation show higher iron and lower nickel contents than that in the original alloy. Both the Debye diffraction patterns on peeled scale after 1 hour and 3 hours at 2000°F in air show the presence of strong spinel type lines with $a=8.36$°A and weaker lines of Fe_2O_3.

The 79 Ni-21 Fe alloy behaves differently from the 50 Ni-50 Fe alloy on oxidation in air at 2000°F in that no abrupt change in the chemical composition of the adherent oxidized layer next to the metallic surface was observed, such as is the case with the latter alloy. Neglecting small variations and also discounting the value obtained after 5 minutes of oxidation due to the very thin oxidized layer, Table VI shows that the ad-

TABLE VI. Relative line intensity of alloys on oxidation.

Alloy	Accumulated time at 2000°F in air	Relative K_α line intensity: Unoxidized specimen Ni/Fe	Unoxidized specimen Cr/Fe	Oxidized specimen Ni/Fe	Oxidized specimen Cr/Fe
17 Cr-Fe	1 min		0.182		0.179
	5 min		0.184		0.200
	30 min		0.171		0.385
(peeled scale)	30 min				0.182
	1 hr		0.172		0.017
	3 hr		0.183		0.006
(peeled scale)	3 hr				0.004
50 Ni-50 Fe	1 min	0.707		0.317	
(blister off)	1 min			0.325	
(scale off)	5 min	0.722		0.129	
	30 min	0.700		0.047	
(peeled scale)	30 min			0.208	
	1 hr	0.697		0.042	
(peeled scale)	1 hr			0.100	
	3 hr	0.715		0.945	
(peeled scale)	3 hr			0.152	
79 Ni-21 Fe	1 min	2.435		2.152	
(scale off)	5 min	2.420		6.520	
	30 min	2.465		4.181	
(peeled scale)	30 min			1.964	
	1 hr	2.401		4.397	
(peeled scale)	1 hr			2.507	
	3 hr	2.380		4.899	
(peeled scale)	3 hr			2.309	
18 Cr-8 Ni Cb stainless steel	1 min	0.157	0.219	0.146	0.238
	5 min	0.160	0.218	0.142	0.273
	30 min	0.138	0.209	0.093	0.254
(peeled scale)	30 min	0.138	0.209	0.220	0.085
(scale off)	1 hr	0.141	0.205	0.356	0.403
	3 hr	0.141	0.221	0.446	0.516
(peeled scale)	3 hr	0.141	0.221	0.368	0.282

[6] Gulbransen, Phelps, and Hickman, Ind. Eng. Chem. 640–652 (1946).

TABLE VII. Diffusion of an alloying element underneath scaled surface.

Specimen preparation	Ratio of intensity of Cr line to Fe line
Scraped clean of loose scale	0.063
1st polish	0.191
2nd polish (most of red scale off)	0.192
3rd polish (almost clean)	0.203
4th polish (clean)	0.318
Unoxidized surface	0.182

herent layer remains richer in nickel content than that in the original alloy despite the repeated peeling-off of the scale. The chemical composition of the peeled scale after various time periods of oxidation, as shown in Table VI, does not seem to deviate far from the original alloy composition. The Debye diffraction pattern on the peeled scale after 1 hour at 2000°F in air shows the presence only of the spinel type lines of $a=8.37$°A. The lattice parameter of the spinel cell does not seem to agree with the value reported for $NiO \cdot Fe_2O_3$ in Gulbransen's[6] paper.

On oxidizing an 18 Cr-8 Ni-Cb stainless steel at 2000°F in air the adherent oxidized layer adjacent to the metallic surface grows steadily richer in chromium as well as nickel despite repeated scaling-off on prolonged oxidation. At first, the chromium or nickel content in the adherent oxidized layer starts at a value lower than that in the original alloy, but after oxidation time longer than 30 minutes it becomes much higher than that in the original alloy, and is still climbing after 3 hours of oxidation. The peeled scales after 30 minutes and 3 hours of oxidation show consistently higher nickel content than that in the original alloy with the chromium content increasing from a value under to a value over the original alloy content after 3 hours of oxidation. The Debye diffraction pattern of the peeled scale after 30 minutes of oxidation shows stronger lines of Fe_2O_3 than those of the spinel type ($a=8.36$°A) while the Debye pattern on peeled scale after 3 hours of oxidation shows stronger lines of spinel type than those of Fe_2O_3.

The above results might possibly serve as a key to the solution of the mechanism of oxidation in practical metallurgical alloys which are subjected to service at elevated temperatures. The data might also be of interest to the ceramic industries in relation to wetting and sealing of glasses, structure of coatings, etc.

The established experimental evidence from the literature and from the present investigation shows that the diffusion of alloying elements plays an all important role in the mechanism of oxidation at elevated temperature. If an alloying element is known to take part in oxide film formation, the picture of diffusion would not be complete without a definite proof by some physical means that a change in the percentage of that alloy element from the average composition of the alloy takes place in the metallic layer directly underneath the oxide film. In the present investigation a specimen of 17 Cr-Fe alloy after oxidation in air at 2000°F for 3 hours was scraped as clean as possible from loose oxide film, and then was gradually polished down in steps toward the center section with fine 00 emery paper. The *K*-alpha lines of Cr and Fe fluorescent radiations were scaled during each step.

The results shown in Table VII indicate the existence of a chromium-rich metallic layer next to unoxidized metal. From that chromium-rich metallic layer to the boundary of the nearest oxide film, the intervening metallic layers gradually impoverish in the chromium content, with the lowest chromium content next to the

TABLE VIII. Analyses of extracts from alloys by x-ray diffraction and fluorescence.

Alloy composition	Method of extraction	Diffraction data	Fluorescent data	Remarks
Slag inclusion from a titanium bearing alloy	As received	$CaTiO_3$, orthorhombic, $a=10.762$°A, $b=7.645$°A, $c=10.886$°A	Figure 3(a)	Chemical analysis: 6.13% CaO, 21.35% TiO_2
18 W-4 Cr-1 V high speed steel	Anodic 10% HCl aqueous	$(Fe, W, V, Cr)_6C$, face-centered cubic, $a=11.04$°A	Figure 3(b)	
19 Cr-9 Ni-W-Mo steel	Anodic $FeCl_3$ aqueous	Sigma (50 Cr-50 Fe), tetragonal, $a=8.790$°A, $c=4.559$°A	Figure 3(c)	
S-816 high temperature alloy 40 Co, 20 Ni, 20 Cr, 4 W, 4 Mo, 4 Cb, 0.4 C, balance Fe	Anodic $FeCl_3$ aqueous	CbC, face-centered-cubic, $a=4.44$°A	Figure 3(d)	
18 Cr-8 Ni-Ti AISI type 321 stainless steel	Bromine	TiC, face-centered-cubic, $a=4.32$°A	Figure 3(e)	
18 Cr-8 Ni-Cb AISI type 347 stainless steel	Bromine	CbC, face-centered-cubic, $a=4.44$°A	Figure 3(f)	

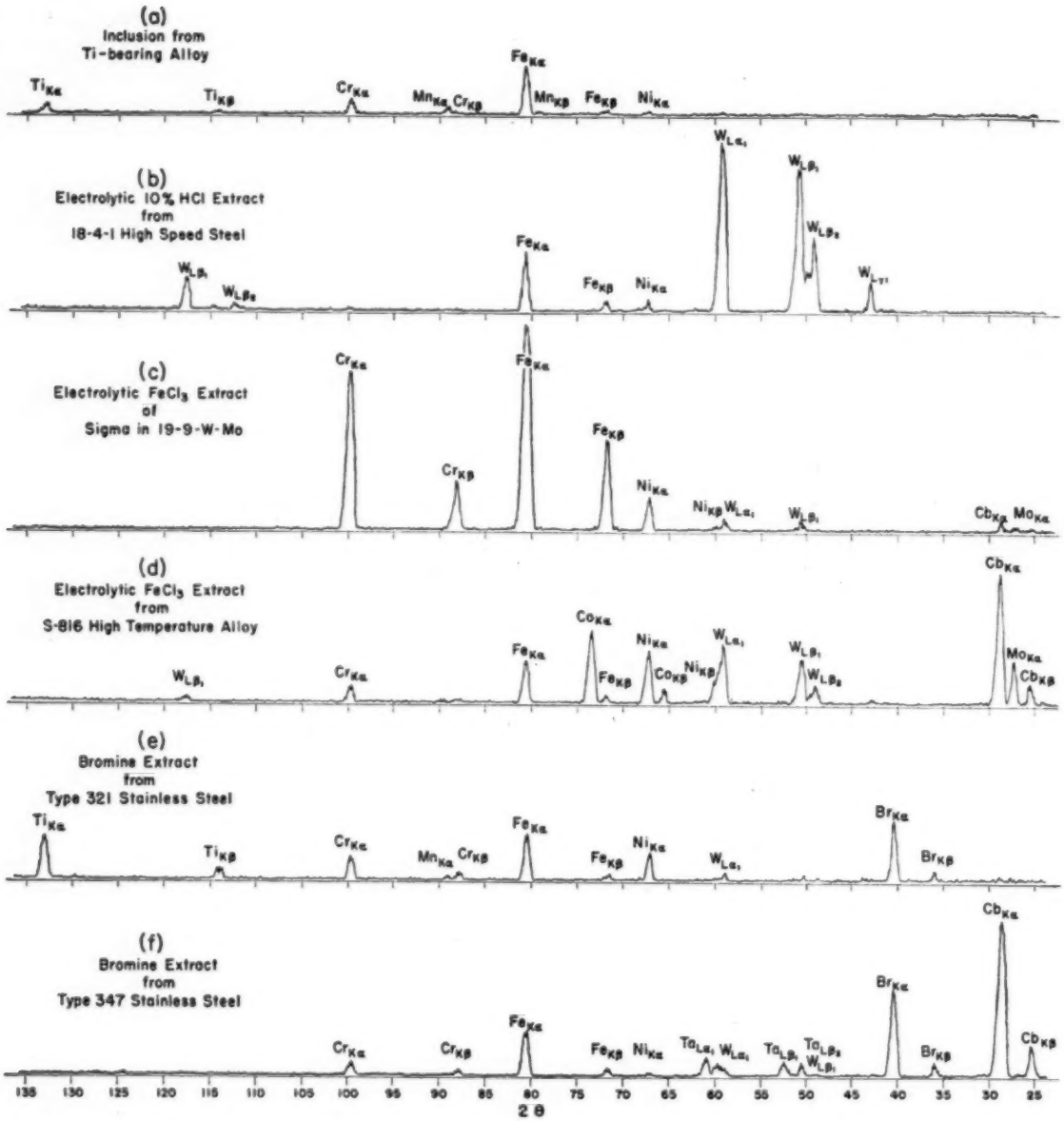


FIG. 3. Fluorescent recordings of powder samples consisting of extracted phases which have been identified by x-ray diffraction as nominally (a) $CaTiO_3$, (b) $(Fe, W, V, Cr)_6C$, (c) sigma, (d) CbC, (e) TiC, (f) CbC.

oxide film boundary, and become much lower than that in the average alloy composition of 17 percent. The reason for the impoverishment of the chromium content in metallic layers immediately underneath the oxide film can be explained as due to the enrichment of chromium in the initial superficial oxide film. The existence of the underlying chromium-rich metallic layer might perhaps be interpreted as due to movement of chromium from the central section outward during oxidation at the elevated temperature.

SUPPLEMENTARY INFORMATION BY FLUORESCENT ANALYSIS ON SUBSTITUTIONAL PHASES

Among the simplest binary systems, wide solid solubility ranges frequently exist among intermetallic compounds of both open-maximum and peritectic types. The problem of solid solubility in intermetallic compounds becomes more complicated in commercial alloys in which added alloying elements and unavoidable impurities from various sources take part. Quite frequently half of an alloying element in an intermetallic compound, usually identified with that element, can be replaced by other similar elements, and yet the intermetallic compound retains the same crystal form of only slightly different dimensions. Therefore, in order to identify a phase completely, diffraction data should be supplemented by chemical analysis. When only a small amount of sample is available, such as is the case with extracts from limited samples, certain corrosion products, etc., x-ray fluorescent analysis is capable of

supplying additional information which helps to define the true status of a phase in an alloy.

The holder used in the present investigation for analysis of powder specimens is made from aluminum and is provided with a one-inch diameter cavity fitted snugly with transparent plastic disks. Layers of powder of various thicknesses can be accommodated by using different combinations of disks. Only the fluorescent radiations from the powder specimen are detected by the Geiger counter, as both aluminum and plastic yield negligibly small background line intensities.

A number of commercial alloys were extracted by chemical and electrolytic methods for carbides and other intermetallic phases. The extracts were analyzed by x-ray diffraction and fluorescence. The results obtained are listed in Table VIII and Fig. 3.

Figure 3(a) shows the presence of extraneous phases in the slag inclusion which do not appear in the diffraction pattern. Figure 3(b) shows the predominant amount of W, large amount of Fe, noticeable amount of Ni, negligible amount of Cr in the so-called $(Fe, W, V, Cr)_6C$ carbide. The amount of V in the carbide is difficult to determine by this method as V gives a very low fluorescent yield. The nickel in the extract is derived from the impurity in the melting scrap. Figure 3(c) shows the solubility of Ni, W, Mo, and possibly Cb in the Fe-Cr sigma phase. Figure 3(d) shows the definite trend of substitution of Cb by W and Mo in columbium carbide. The fluorescent recording indicates some matrix contamination, although the diffraction pattern shows carbide lines only. Bromine chemical extraction was adopted for carbides shown in Fig. 3(e) and Fig. 3(f) because this method produces extracts whose diffraction patterns show no contamination of the austenitic matrix, while the electrolytic extraction methods applied to these steels cause intergranular disintegration with matrix contamination as a result. The fluorescent analysis on these carbides, however, shows the serious contamination of bromides which do not appear in the diffraction patterns. No conclusion can be drawn on the presence of Fe, Cr, Ni, and W in the extract because of their possible identities as bromides or matrix. Figure 3(f) shows significantly the existence of Ta in the columbium carbide. It is generally known that Ta coexists with Cb in the ores, the separation of the two being quite difficult.

Deformation and Fracturing of Thick-Walled Steel Cylinders under Explosive Attack*

JOHN PEARSON AND JOHN S. RINEHART
Michelson Laboratory, U. S. Naval Ordnance Test Station, Inyokern, China Lake, California
(Received October 17, 1951)

Tests have been conducted on annealed heavy-walled cylinders of low carbon (1020) steel internally loaded by explosive charges. The purpose of these tests was to obtain basic information on the manner and type of fracturing and plastic flow obtained by extremely high pressures acting for short durations. It was observed that all of the cylinders tended to fracture in long fragments in the same basic pattern, but with variations because of wall thickness and manner of loading. Experimental results indicate that the radial cleavage type of fracture is initiated within the cylinder wall and propagated to the surface, while the shear type fracture associated with the inner portion of the cylinder wall appears as an independent energy relieving process.

Microstructure analysis of the cylinder fragments shows a definite relationship between the type of fracture and the amount of distortion of the grain boundary. Considerable shock twinning is present in the cylinder fragments, and the grain structure near the inner surface shows severe distortion and flow. Strain measurements indicate that considerable plastic strain occurs during the explosive loading and cleavage fractures were always observed to occur in association with considerable plastic deformation.

INTRODUCTION

THE purpose of this investigation was to study the nature of the deformation and fracturing produced in thick-walled steel cylinders by extremely high internal transient pressures. A peak pressure in the neighborhood of 4,000,000 psi[1,2] is reached through the use of high explosive charges. The duration of the pressure is not known precisely, but it is believed to be of the order of a few milliseconds.

The problem of the internally loaded thick-walled cylinder has been extensively studied theoretically[3–9] for both the elastic and plastic cases, and a number of investigators have reported experimental observations regarding the failure of cylinders under more gradually applied internal loads. There are, however, few data available on the effects of explosive loading on such a cylinder. Explosive loading introduces extremes in the magnitude and time rate of loading which are not present in the more generally used types of tests. A complete theoretical treatment of the explosively loaded cylinder would necessarily include a consideration of the manner in which the transient stress waves are propagated through the cylinder. Such a treatment has not been published. It is anticipated that the present investigation will throw considerable light on the nature of the problem. It was felt that the plastic solutions that are available, as well as extensions of the elastic solution for thick-walled cylinders, could not be used to predict the behavior of cylinders under internal explosive attack. Furthermore, as Bridgman[10] has remarked, the phenomena of fracture are complicated and obscure. For this reason, workers in the field have been encouraged to acquire the broadest possible knowledge of the phenomena. It was partially in line with this thought that the present tests were conducted.

TABLE I. Test cylinder data.[a]

Wall thickness (inches)	Steel analysis no.	External grids only used	No grid systems used	Number of cylinders fired
½	1		×	3
¾	1		×	3
1	1		×	3
1½	1		×	3
¾	2	×		1
1	2	×		1
1½	2	×		1
2 3/16	3	×		1
2 7/16	3	×		1
4[b]	4	×		2

[a] *Note*: All cylinders have 1-in. i.d.
[b] This cylinder 6 in. long, all other cylinders 8 in. long.

EXPERIMENTAL WORK

All tests were conducted with the cylinders mounted in a vertical position on several layers of Celotex as shown in Fig. 1. The cylinders were fired in shallow pits in sandy soil with the outer surface of the cylinder free from any restraint. In all cases, Composition C–3 was the explosive. It was detonated with an Engineer Special Cap that was inserted in the upper end.

The inside diameter of each cylinder was the same, 1 in. Wall thicknesses ranged from ½ in. to 4 in. All cylinders were made 8 in. long, except the one with a 4-in. wall whose length was limited to 6 in. by the

* Work supported by Armament Branch, ONR.
[1] Evans, James, and Pack, Proc. Phys. Soc. (London) **60**, 1 (1948).
[2] C. S. Robinson, *Explosions, Their Anatomy and Destructiveness* (McGraw-Hill Book Company, Inc., New York, 1944).
[3] S. Timoshenko, *Theory of Elasticity* (McGraw-Hill Book Company, Inc., New York, 1934).
[4] A. Nadai, *Theory of Flow and Fracture of Solids* (McGraw-Hill Book Company, Inc., New York, 1950), second edition, Vol. 1.
[5] MacGregor, Coffin, Jr., and Fisher, J. Franklin Inst. **245**, 135 (1948).
[6] MacGregor, Coffin, Jr., and Fisher, J. Appl. Phys. **19**, 291 (1948).
[7] P. W. Bridgman, J. Appl. Phys. **19**, 302 (1948).
[8] P. G. Hodge, Jr., and G. N. White, Jr., J. Appl. Mech. **17**, 180 (1950).
[9] Hill, Lee, and Tupper, Proc. Roy. Soc. (London), Series A **191**, 278 (1947).
[10] P. W. Bridgman, "Fracture and hydrostatic pressure," in *Fracturing of Metals* (Am. Soc. Metals, 1948), p. 246.

amount of material available. The several cylinders used in the tests are listed in Table I.

The cylinders were machined from solid round bars of 1020 steel, rather than from tubing in an effort to obtain homogeneity and uniformity of material in the cylinder walls. They were then annealed in a reducing atmosphere at 1600°F. The cylinders with outside diameters of 4 in. or less were machined from bars $4\frac{1}{4}$ in. in diameter. The other cylinders were machined from bars of 6 and $9\frac{1}{4}$ in. in diameter. Complete analyses of the several stocks used are given in Table II.

In a few cases fine $\frac{1}{4}$-in. rectangular grid systems were scribed on the outer surface of the cylinder in order to measure permanent strain. The depth of the scribed line was never greater than 0.003 in. The cylinders were annealed after the grids were put on in order to relieve any local stresses caused by the scribing process. The use of finely scribed grid lines on the interior surface did not prove feasible as the longitudinal grids became lines of fracture.

A few minor problems were encountered in the recovery of the fragments. Complete recovery of all fragments was difficult. However, the percentage of recovery was sufficiently high (up to about 95 percent by weight) to permit a comprehensive comparison study of the fragments from the various cylinders tested. This study indicated a high degree of reproducibility of the tests. A few fragments were damaged by striking objects. But here again, the number was not sufficient to interfere in any way with the experimental studies which were conducted.

The recovered fragments were carefully studied. Deformation of the previously mentioned grid system was used as a measure of the amount of permanent straining. Typical fragments were cross-sectioned in order to observe the nature and extent of fracturing. Many of the sectioned fragments were used for studies of changes in microstructure. Each of these studies is discussed in one of the following sections.

GENERAL BEHAVIOR OF CYLINDERS

When a detonating explosive in a thick-walled cylinder is fired at one end, the detonation wave moves down the cylinder and the explosive turns into gaseous products which exert extremely high pressures on the inside of the cylinder. For cylinders 8 in. in length, such as were used in these tests, the time required to propagate this detonation wave the length of the explosive is approximately 27 microseconds. These gaseous products promptly begin to bleed from the ends of the cylinder.

Under these high pressures the cylinders expand radially and then fracture into long thin fragments by a combination of shear and cleavage type fractures. The major fragments obtained also contain many longitudinal breaks which tend to subdivide these pieces and thus increase the number of fragments. For cylinders of the same size and under the same conditions of loading, the extent of this fracturing is about equal so that the number of major fragments obtained is approximately the same in each case. As the wall thickness is increased, the extent of fracturing decreases, and the number of fragments is reduced until a sufficiently thick wall is reached where some plastic expansion occurs but fragmentation does not.

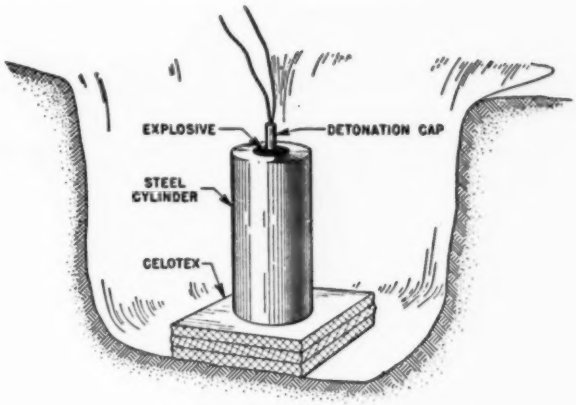


FIG. 1. Firing arrangement for cylinders.

A high degree of reproducibility is obtained when cylinders of the same dimensions are loaded and detonated in the same manner. The type of fractures obtained and the manner in which the cylinders fracture are the same. Also, the size and number of the fragments obtained are quite similar. When these fragments are cross sectioned and etched, microstructure studies show that the physical change in the microstructure of the material is quite reproducible by this type of test.

FRACTURES

All cylinders with a wall thickness of $2\frac{7}{16}$ in. or less broke up into several fragments while the 4-in. wall cylinders remained intact. The $\frac{1}{2}$-in. to $2\frac{7}{16}$-in. thick-walled cylinders fractured in a characteristically similar manner. A photograph of the fragments that were obtained from an annealed cylinder whose wall thickness was $2\frac{7}{16}$ in. is reproduced in Fig. 2. The uniformity in size and the manner in which the major fragments tend to break into smaller fragments are typical of cylinders of other wall thicknesses. The force of the explosion was not sufficient to cause fractures to extend through the entire wall of the 9-in. o.d. cylinder. The effect of the explosive can be seen in Fig. 3 where a fired 9-in. o.d. cylinder is shown next to a similar unfired one. The outer surface of the fired cylinder

TABLE II. Composition of material.

Steel analysis no.	C	Mn	Si	S	P	Cr	Ni
1	0.16	0.67	0.05	0.015	0.007	Trace	Trace
2	0.14	0.66	0.04	0.016	0.007	Trace	Trace
3	0.21	0.96	0.28	0.015	0.017	Trace	Trace
4	0.21	0.95	0.32	0.017	0.010	Trace	Trace

FIG. 2. Fragments from $2\frac{7}{16}$-in. wall cylinder.

showed no signs of fracturing although the inside diameter was increased by as much as 77 percent.

Two basic types of fracture have been observed: shear fracture at approximately 45° to the circumference of the cylinder, and a cleavage fracture extending radially through the cylinder wall. These types of fractures are similar to those observed by N. F. Mott[11] in his work on the fragmentation of shell cases. Both types are evident in Fig. 4 which is a photograph of the cross section of a fragment obtained from a $2\frac{7}{16}$-in. wall cylinder. The interior surface exhibits only shear fractures, while the remainder of the cylinder wall shows the occurrence of cleavage fracturing. Many internal fractures also exist which do not extend to either surface. Such fractures were not described by Mott in his paper.

It may be of interest to discuss briefly the cleavage type of fracture. Frequent references to it in literature tend to be confusing as it is discussed under a wide range of terminology.[4–14] It is often referred to as a cleavage fracture, fibrous fracture, cohesive-cleavage type failure, brittle fracture, granular fracture, crystalline break, and others. Nadai[4] has adequately expressed this type of fracture as one "caused by separation in the grains of polycrystalline material along certain crystallographic planes when the normal stress in these planes reaches a definite mean value, particularly in the lower range of temperatures and at the higher rates of stressing···."

A cleavage fracture is not necessarily a true brittle fracture but may have considerable plastic flow associated with it. In the present experiments with explosive loading of thick-walled steel cylinders, this type of fracture was consistently present with varying degrees of plastic deformation. Davis and Parker[14] in their experiments with thin-walled tubes tested at −138°F with internal pressure only, found that the tubes exhibited up to 50 percent plastic extension before fracture occurred by cleavage. In his paper, Mott[11] also stated that for cylindrical shell cases composed of a ductile material, such as steel, considerable plastic expansion, as much as 50 percent, occurred before the case broke with shear and cleavage type fractures. The amount of permanent strain which occurred in the present case is discussed in a later section.

When the 9-in. o.d. cylinders were cross-sectioned and etched, four distinct regions were noted in the material (see Fig. 5). At the inner surface was a region containing numerous shear fractures which extended to a depth of about $\frac{1}{4}$ in. into the cylinder wall. Starting at a depth of about $1\frac{1}{8}$ in. from the inside surface was another region containing a considerable number of completely internal cleavage fractures which extended radially in the wall of the cylinder. Between the shear and cleavage fractures was a region of unfractured material, and between the internal fractures and the outer surface lay another region of unfractured material.

The existence of internal fractures indicates the probability that when thick-walled cylinders fracture from internal explosive loading, the cleavage fractures initiate within the walls of the cylinders and then propagate in a complex, mutually stress-relieving manner to the inner and outer surfaces. The large unfractured region between the shear and cleavage fractures indicates that the two types of fractures may be initiated as two completely separate systems.

The occurrence of only shear fractures at the inner surface of the cylinders is in agreement with the work of Bridgman[15] where it is shown that tensile fractures occur with difficulty under very high hydrostatic pressures, and that at high pressures the break becomes a shearing fracture rather than a tensile break. He has also shown[16] why it is unlikely that rupture would start at the inner surface under these loading conditions since such a rupture at the inner surface would not result in a decrease of the total potential energy of strain in the metal, as the internal pressure would be acting in a direction to oppose the release of volume compression resulting from such a rupture. On this basis, it appears likely that under explosive conditions of loading a cylindrical region extending from the inner

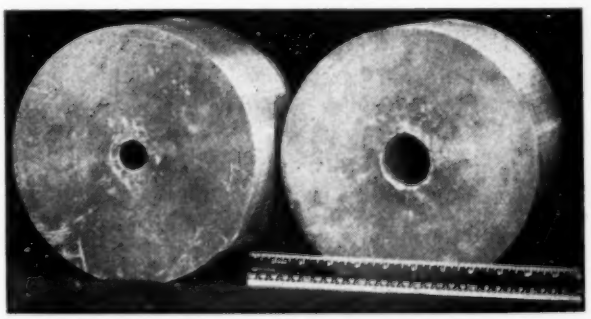


FIG. 3. Four-inch wall cylinders before and after internal explosive loading.

[11] N. F. Mott, Proc. Roy. Soc. (London) Series A **189**, 300 (1947).
[12] E. A. Davis, J. Appl. Mech. **15**, 216 (1948).
[13] S. J. Fraenkel, J. Appl. Mech. **15**, 193 (1948).
[14] H. E. Davis and E. R. Parker, J. Appl. Mech. **15**, 201 (1948).
[15] P. W. Bridgman, Revs. Modern Phys. **17**, 3 (1945).
[16] P. W. Bridgman, *The Physics of High Pressure* (G. Bell and Sons, London, 1949).

surface approaches the conditions of hydrostatic pressure. The cleavage fractures are initiated at the outer edge of this hydrostatic region. The shear fractures then appear as an energy relieving process in the region of greatest deformation after the tube has expanded plastically and the cleavage fractures have been initiated within the wall.

An interesting observation was made regarding the shear fractures in the 4-in. wall cylinders. The shear fractures all propagated in the same direction for about 90° of the inner surface at which time they reversed their direction of propagation for the next 90°. For each cross section studied, this reversal occurred four times and roughly corresponded to four quadrants of a circle. Occasionally, at the point where two quadrants met, two shear cracks propagated at right angles to each other from the same point.

The foregoing observations of the manner in which thick-walled cylinders fracture under internal explosive loading indicate that extensions of elastic theory to fracture initiation are not applicable in the present case. The application of plastic theory is complicated by the introduction of several factors resulting from the extreme pressures and time rate of loading used. The results of the present tests, namely, that the location of tensile fractures are initiated within the cylinder wall, show that the location of fracture initiation in explosively loaded thick-walled cylinders is different from the observed results for cylinders loaded under high internal pressures but at slow rates of straining. In the latter case the fractures were observed to begin at the outer surface of the cylinder.

The mathematical solution for the static loading of thick-walled cylinders within the elastic range is well known. The stress equations for such a cylinder loaded by an internal pressure, where it is assumed that a plane transverse section of the cylinder remains plane and parallel to itself, have been fully developed and can be written in the form[3]

$$\sigma_r = \frac{a^2 p}{b^2 - a^2}\left(1 - \frac{b^2}{r^2}\right)$$

$$\sigma_\theta = \frac{a^2 p}{b^2 - a^2}\left(1 + \frac{b^2}{r^2}\right),$$

where a and b denote the inner and outer radii of the cylinder, respectively, and p denotes the uniform internal pressure. The tangential stress reaches a maximum value at the inner surface of the cylinder where

$$(\sigma_\theta)_{\max} = \frac{p(a^2+b^2)}{b^2-a^2}.$$

Therefore, the maximum value of σ_θ is always numerically greater than the internal pressure and approaches this quantity as material is added to the outside wall of the cylinder. The values for shearing stress and tangential strain also are a maximum at the inner surface where

$$\tau_{(\max)} = (\sigma_\theta - \sigma_r)/2 = pb^2/(b^2-a^2)$$

$$\epsilon_{(\max)} = (\sigma_\theta - \mu\sigma_r)/E.$$

It was realized that none of the above criteria would be realistic in the present case since a certain amount of plastic flow must be expected, and under such conditions the relations between stress and strain which were assumed in deriving the above formulas no longer hold. However, it might be supposed that elastic theory could be extended in a very general manner to indicate that the fracturing of a thick-walled cylinder subjected to high internal pressure might be expected to start at the inner surface. The inadequacy of extending

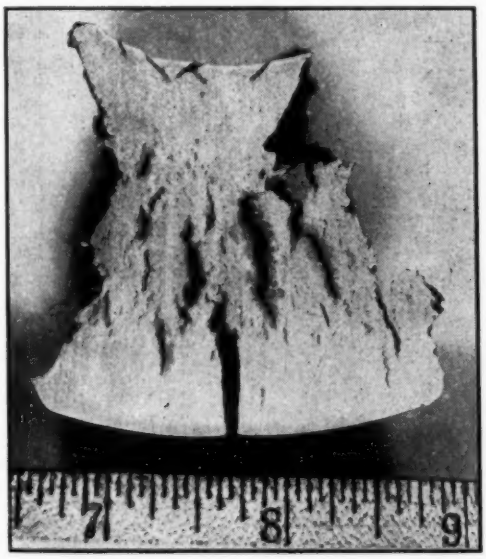


FIG. 4. Fractures in cross section of thick-walled cylinder fragment.

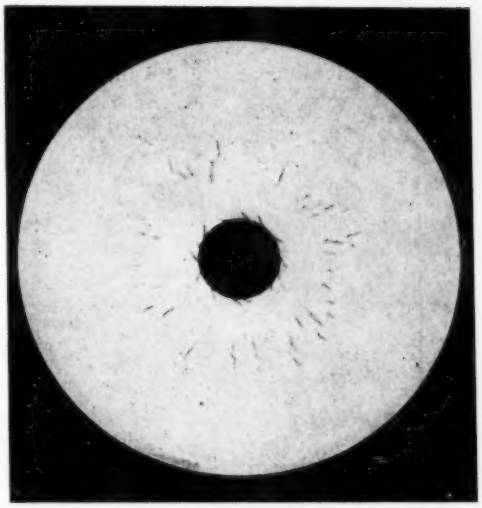


FIG. 5. Cross section of 4-in. wall cylinder showing effects of internal explosive loading.

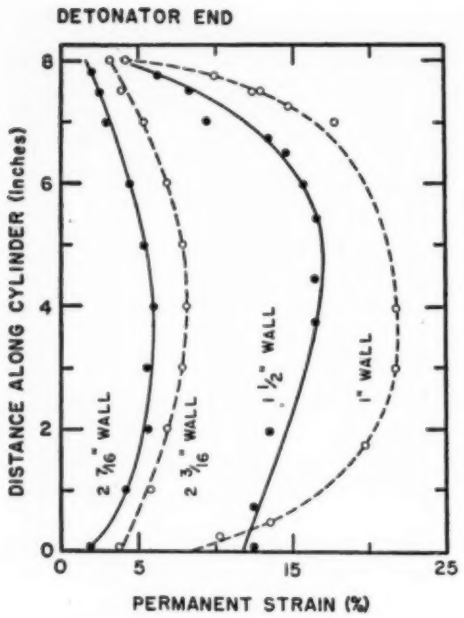


FIG. 6. Circumferential permanent strain on outer surface of cylinder wall.

elastic theory even in this generalized manner to explosively loaded thick-walled cylinders is shown by the fact that the major fractures did not start at the inner surface as the above criteria would indicate.

Extensive theoretical treatments have been given for thick-walled cylinders in the plastic-elastic and the fully plastic states[4–9] for static loading and fairly low rates of strain. These treatments have included discussions of the plastic yielding of thick-walled cylindrical tubes acted upon by various combinations of internal pressure, external pressure, and end loads based upon different infinitesimal strain and finite strain conditions. There is not yet complete agreement, however, on a generalized treatment that fully covers all conditions of loading.

Application of plastic theory to thick-walled steel cylinders loaded with explosive charges is complicated by several factors resulting from the conditions of loading. At very high rates of loading the initial yield pressure is probably independent of the wall thickness since the high pressures are applied so rapidly that yielding at the interior surface may start before the outer surface is aware of loading. Other factors that must be considered are a delay in plastic initiation,[17] an increase in ultimate strength with rate of loading,[18] the fact that resistance to fracture may depend on the state of stress,[15,19] and the effect of strain hardening. These effects make the problem very complex and the obtaining of a complete solution to the problem difficult.

[17] D. S. Wood and D. S. Clark, Trans. Am. Soc. Metals **43**, 571 (1951).
[18] D. S. Clark and D. S. Wood, Trans. Am. Soc. Metals **42**, 45 (1950).
[19] J. S. Rinehart, J. Appl. Phys. **22**, 55 (1951).

The results of the present tests indicate that the experimental observations of other investigators regarding the manner of failure of internally loaded cylinders under relatively slow rates of loading cannot be extended to include the explosive loading of such cylinders. The high pressure work of Bridgman[16] indicates that the rupture of cylinders made of ordinary grades of steel invariably starts at the outside surface and travels inward, rather than starting at the inside and running out. Another manner of fracturing has been proposed by Davis[12] who has expressed the view that where the fracture is completely brittle† the fracturing may occur at both surfaces and spread toward the central part of the wall and was suggested as a means of explaining the herringbone type of break. However, it was believed that the extremely short duration of loading and the magnitude of the pressure obtained by using explosive charges might give results entirely different from those expressed in the above works. In fact, the authors found considerable evidence that the major fractures caused by internal explosive loading were initiated within the walls of the cylinders rather than at either surface. Davis has also discussed the possibility of associating shear and cleavage fracturing in hollow cylinders to the speed of crack propagation in connection with tests of his in which tubular steel specimens with closed and open ends were stressed by internal fluid pressure.

STRAIN MEASUREMENT

Finely scribed rectangular grid systems were placed on the outside of the six thickest walled cylinders to measure permanent strain. Internal grids were tried but abandoned since the grid lines affected the shear

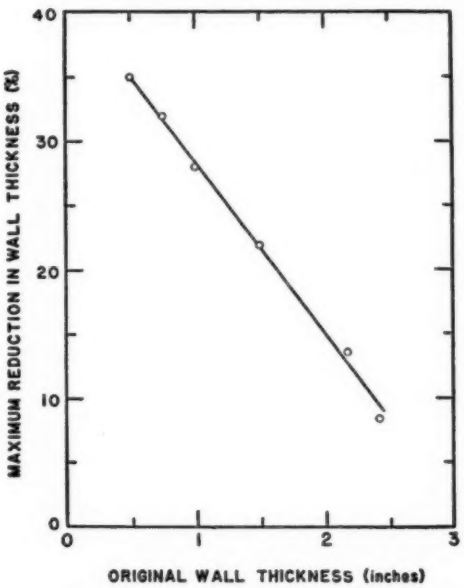


FIG. 7. Maximum reduction in wall thickness of fragmented cylinders.

† See prior discussion in paper regarding brittle fracturing.

fractures at the interior surface and gave rise to the formation of long wedge-shaped fragments. Neither external nor internal grid systems were used on the ½-in. wall cylinders as the scribed lines interfered with the method of fracture.

Considerable plastic deformation occurred in the radial and tangential directions. The amount of circumferential deformation varied with position along the cylinder and with wall thickness. Curves characteristic of the circumferential permanent strain on the outer surface of several cylinders are given in Fig. 6. The maximum strain for a 1-in. wall cylinder is 21.6 percent and that for a $2\frac{7}{16}$-in. wall cylinder, 6 percent. The curves indicate that the maximum circumferential strain increases as the wall thickness decreases. The curves also show that for an individual cylinder the permanent strain is the smallest near the ends of the cylinder and reaches a maximum not far from the middle. The maximum strain for a 1-in. wall cylinder, for instance, reaches 21.6 percent near the center, while it drops to a minimum of 4.2 percent at the upper end. Measurements of the 4-in. wall cylinders showed an increase of from 58 to 77 percent for the inside diameter over the original 1-in. dimension, while the outside diameter showed a maximum increase of 0.92 percent over the original 9-in. dimension.

As the wall thickness was reduced, the relative reduction in the wall resulting from the plastic expansion increased. This value varied along the fragment. The greatest percent reduction occurred about halfway down the cylinder in the region of maximum circumferential strain. The maximum relative reduction in the walls of six cylinders is shown in Fig. 7. It is of interest to note that from a wall thickness of ½-in. up the curve has a linear relationship.

Longitudinal strain up to 6 percent was evident in the central portion of fragments from the 1 and 1½-in. wall cylinders. This strain was apparently because of the severe buckling which many of these fragments underwent. The buckling is clearly evident in Fig. 8 which shows a fragment from a 1½-in. wall cylinder. No elongation was evident in portions of the fragment which had remained plane. Fragments from the ¾-in. wall cylinders, which remained essentially flat with only small amounts of buckling present, showed maximum amounts of permanent longitudinal strain of about 2½ percent. The flat portions of these fragments showed no strain in the longitudinal direction.

FIG. 8. Fragment from 1½-in. wall cylinder.

FIG. 9. Shear fracture caused by explosive loading.

MICROSTRUCTURE ANALYSIS

The microstructure of the cylinder fragments was characterized by severe grain distortion and shock twinning. The amounts of twinning and the degree of grain distortion were both greatest near the inner surface and both decreased as the outer surface was approached. It is likely that considerable slip occurred during the expansion process but was not evident since the specimens were polished and etched.

The depth of the shear fractures from the inner surface closely corresponded to the depth of the heavy plastic flow region where the grains were elongated in the circumferential direction. The shear fracture ended, or changed to the cleavage type fracture, in the region where the grain structure showed no further flow or boundary elongation. The relation between the grain distortion and shear fractures is shown in Fig. 9, which is a photomicrograph taken of an area $\frac{11}{64}$ in. from the inside surface. The photograph also illustrates another point of interest, namely, that many of the shear fractures do not propagate as complete fractures, but rather as a series of breaks of varying lengths with each break interposed by unbroken material.

Considerable shock twinning has been observed throughout the cylinder wall where full fragmentation occurred, but in the 4-in. wall cylinders this shock

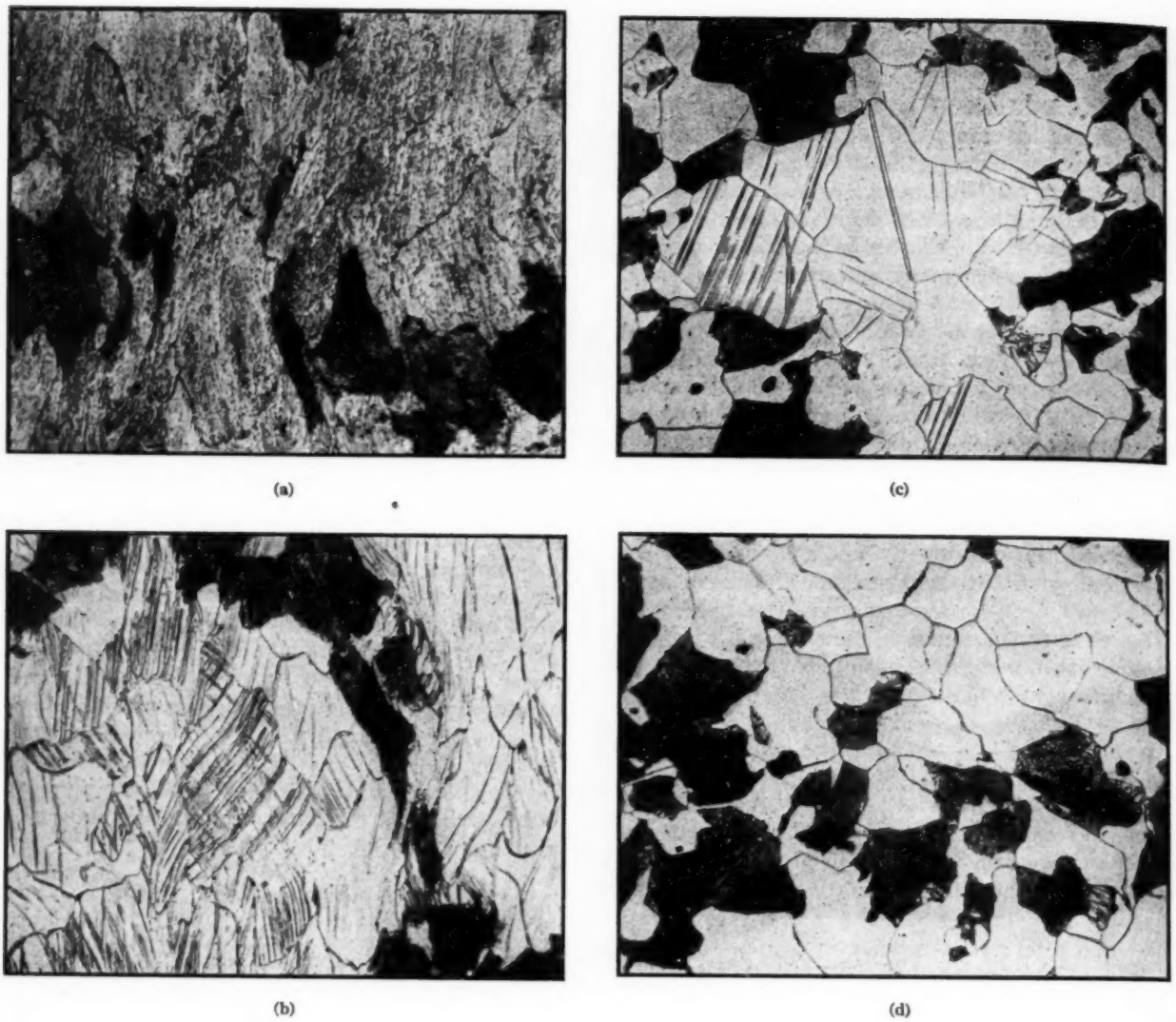


FIG. 10. Microstructure in fired 4-in. wall cylinder. (a) $\frac{1}{64}$ in. from explosive. (b) $\frac{5}{64}$ in. from explosive. (c) $1\frac{7}{16}$ in. from explosive. (d) $\frac{5}{32}$ in. from outer surface.

twinning ceased near the outer region where the internal fracturing also ended. Extensive study has been done on the phenomena of Neumann bands or shock twins. It is known[20] that a rapid loading of the material is one cause of shock twinning. The distortion of the shock twins near the inside surface of the cylinder gives an indication of the manner in which the twins occurred under the dynamic loading, and of how the grains were then deformed, possibly after a slight delay period.

The change in the microstructure of the steel progressing from the inner to the outer surfaces of one of the 4-in. wall cylinders is shown in Fig. 10. Figure 10(a) shows the heavy flow associated with the metal nearest the explosive and was taken at $\frac{1}{64}$ in. from the inner surface. Figure 10(b), taken at $\frac{5}{64}$ in. from the inside surface, still shows the presence of some grain distortion and also the heavy shock twinning which is characteristic of this region. The third picture, Fig. 10(c), shows that at a distance of $1\frac{7}{16}$ in. from the inner surface the grain elongation has ceased although the shock twins occur to a lesser extent. Figure 10(d) shows the microstructure at a distance of $\frac{5}{32}$ in. from the outer surface in a region where no fracturing has occurred. The microstructure of Fig. 10(d) appears the same as the original grain structure of the material.

[20] Smith, Dee, and Young, Proc. Roy. Soc. (London) Series A 121, 477 (1928).

CONCLUSIONS

1. Cleavage fractures were always observed to occur in association with considerable plastic deformation.

2. Two basic types of fracture were obtained; shear fracture and cleavage fracture. Microstructure study of the fragments indicated a definite relationship between the type of fracture and the amount of distortion of the grain boundary.

3. Considerable shock twinning was present with each fracture, and the intensity of the Neumann bands decreased with increasing distance from the interior wall.

4. There is considerable evidence that the fractures initiate within the wall and then propagate in a mutually stress relieving manner to the surfaces. It is thought that the shear fractures present at the interior surface are the result of an energy relieving process which occurs after the initial high pressure within the cylinder has been released and after the cylinder itself has undergone plastic expansion.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Mr. J. Craig McLaughlin for handling the ordnance work and to Mr. Walter M. Winfrey, Jr., for the microphotographs.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 4 APRIL, 1952

Influence of Cold-Work on the Hall Effect of a Metal

BERNARD WIENER AND GERHART GROETZINGER
Lewis Laboratory, National Advisory Committee for Aeronautics, Cleveland, Ohio
(Received October 17, 1951)

Cold-work in copper to the extent of 75 percent reduction in the cross-sectional area resulted in an increase in the Hall constant of 0.01 ± 0.002 times the Hall constant for annealed copper. The value obtained for the Hall constant for annealed copper was minus 5.33×10^{-12} volt-cm/ampere-gauss.

INTRODUCTION

IT has been known for a long time that the electrical conductivity of a pure metal is affected by cold-work. In copper at room temperature,* e.g., the conductivity decreases approximately 2 to 2.5 percent for heavy cold-work.[1,2]

According to the free electron theory of metals[3] the electrical properties can be represented in terms of two quantities. Usually, the following are chosen: the number of free electrons per cubic centimeter, n; and the mean free path of the electrons, l. In terms of these quantities, the electric conductivity σ can be represented by the following equation:

$$\sigma=e^2nl/hk, \tag{1}$$

where e is the charge of the electron and, k, the wave number corresponding to the surface of the Fermi distribution. The quantity h has the usual meaning.

In terms of this theory a change of the resistivity indicates, therefore, a change in the product nl. From a measurement of the conductivity no knowledge about the influence of the cold-work on n or l individually can be obtained. So far it has been (at least tacitly) assumed that the cold-work influences the mean free path rather than the number of free electrons, which is for the case of copper taken as being equal to the number of atoms.

In order to determine both n and l, a second experiment is therefore necessary which yields these quantities in a different functional relation than the product. The Hall effect can be used to great advantage for this purpose since its magnitude depends simply on n but not on l. As a measure of this effect the Hall constant R is generally used; this quantity is defined in the following way: If the direction of the magnetic field H is perpendicular to the current density j of the primary current i, the Hall constant is equal to the component E_H of the electric field perpendicular to H and j divided by the product of H and j, or in symbols,

$$R=E_H/jH, \tag{2}$$

the sign of R being defined as positive if E_H has a direction corresponding to a deflection which free positively charged particles carrying the primary current i would suffer under the influence of the magnetic field, H.

Measurements of the Hall effect are usually performed on a conductor of rectangular cross section (in form of a strip) where the width w is chosen very much larger than the thickness d and the magnetic field is perpendicular to the strip. The Hall constant R is then equal to

$$R=Vd/Hi, \tag{3}$$

where i is the electric current flowing through the bar, and $V=E_Hw$ is the voltage resulting from the Hall effect.

According to the above-mentioned theory, the constant R can be represented by the following equation:

$$R=1/ne. \tag{4}$$

In order to shed more light on the mechanism of cold-work it was decided to investigate the magnitude of the Hall effect as a function of cold-work. As a material,

[1] Kinnosuke Takahasi, Science Reports of Tohoku University, 19, 265 (1930).
[2] J. S. Koehler, Phys. Rev. 75, 106 (1949).
[3] See, e.g., A. Sommerfeld and H. Bethe, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1933) Vol. 24, part 2, p. 333.

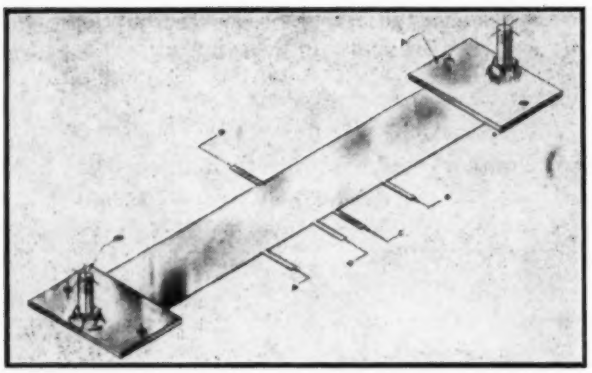


FIG. 1. Sample strip for measuring the Hall effect.

copper was used. Experiments of this type have never been performed before, possibly in view of the fact that a change in the effect resulting from cold-work must necessarily be rather small (smaller than the change of resistivity) if there is a change at all.[4] Besides the measurement of the Hall effect, the change of conductivity resulting from cold-work was investigated for all the samples in which the Hall effect was measured.

EXPERIMENTAL PROCEDURE

Since, as mentioned before, the change in the Hall effect by cold-work can hardly be expected to exceed the change of two percent which occurs in the case of the conductivity, considerable time was spent to develop a procedure with the rather small electromagnet at our disposal for detecting changes as small as one-half of one percent after annealing of the previously cold-worked sample. To achieve this accuracy it was necessary to repeat measurements on a sample in each state about twenty times and to carry out the experiment on many samples. The experiment was conducted as follows: Each sample consisted of a copper strip of approximately 0.014-cm thickness, 5-cm width, and 25-cm length and was kept sandwiched between Bakelite plates in a cooled oil bath during the measurement of its Hall constant and its resistance. The strips were cut from cold-worked sheets in the form shown in Fig. 1 so that the protruding terminals a to e, which were used for measurement of the Hall voltage and of the resistance, comprised an integral part of the strip material. Copper wires were soft soldered to these tails and the junctions immersed in the oil bath. Keeping the junctions in this way separated from the part of the strip carrying the heavy primary current supplied at the terminals A and B prevented disturbance of the measurements by thermoelectric forces resulting at the terminals a to e. During the Hall effect experiments it was not necessary to register the temperature of the oil bath since the Hall constant in copper is known to be almost independent of temperature at temperatures around 20 to 30 degrees centigrade.[5,6] The purpose of the bath was in this case to keep the junctions a, b, and c at the same temperature. No special attempt was made to obtain the condition for the isothermal or adiabatic Hall effect since the emphasis in this investigation was on the change in the Hall effect due to cold-work and not its absolute value under a particular condition. The method used was similar to that proposed by F. Koláček[7] in which the potential of a point on one side of the strip (terminal a on Fig. 1) is matched (in the absence of the magnetic field) by the potential at the tap of a voltage divider connected between two points on the other side in our case, b and c. This method is preferable to applying the voltage divider across points A and B since, as already found by Novák and Macků,[8] in this case the Hall emf is reduced by a factor of two.

In the course of the experiments the arrangement was changed and improved in many respects. If not expressly mentioned the following refers to the "final experiments."

The circuit used for measuring the Hall emf is shown in Fig. 2. The primary current was read on a 50-mv Weston millivoltmeter in connection with a 200-ampere shunt. During the measurement the current was kept as close to 200 amperes as possible, the millivoltmeter never deviating from this reading by a deflection corresponding to more than one-half of an ampere. The Hall voltage was measured by compensation: As a null indicator, a type H. S. Leeds and Northrup reflecting galvanometer of an internal resistance of 22.4 ohms, a critical damping resistance of 42 ohms, and a sensitivity of 0.09 μV/mm with the scale at one-meter distance was used. The compensating voltage was obtained from a specially adapted Lewis Engineering Company potentiometer (shown inside of dashed rectangle in Fig. 2) in connection with a voltage divider consisting of the resistances R_1 and R_2 of 42 ohms and 4500 ohms, respectively. The value chosen for R_1 made the total resistance of the galvanometer circuit approximately equal to the critical damping resistance. In order to ascertain the zero point of the galvanometer, a Leeds and Northrup "Pinch Type" double pole switch with copper contacts (switch No. 1 in the diagram) was provided which made it possible to remove the galvanometer from the circuit and short it through the resistance R_3.

The current between the terminals A and B was obtained from a constant current (welding) generator and the magnet current from a motor generator set. In

[4] However, there is an investigation (W. F. Leverton and A. J. Dekker, Phys. Rev. **80**, 732 (1950)) in which the Hall effect is measured for a thin metal film of antimony which was obtained by distillation. It was observed that heating this film to a temperature of 510°C increased the Hall effect by 13 percent. The transition of such a film during the heating is, of course, an entirely different process than the change of a material from the cold-work to the annealed state.

[5] H. Alterthum, Ann. Physik, **38.5**, 933 (1912).

[6] E. H. Hall, Proc. Am. Acad. Sci. **72.9**, 301 (April, 1938).

[7] *Theoretischer Beitrag zu den Messmethoden des Halleffektes* (Electrina Magnetismus, Prague, 1904), p. 246.

[8] N. Novák and B. Macků, Čas. Math. Fys. **38**, 47 (1908).

order to obtain an absolute value of R, measurements of the magnetic field were performed at various magnet currents with a General Electric indicating fluxmeter.

Measurements were performed in the following way. With a current between A and B but the magnet unexcited, the resistances R_4 and R_5 were adjusted so that no deflection of the galvanometer resulted. Then, the magnet was excited and the voltage necessary for compensating the Hall voltage recorded. Subsequently, the direction of the magnet current was reversed and the Hall voltage compensated again with the compensating voltage reversed by the use of the switch No. 2 in the diagram. About ten such readings were obtained for each direction of the magnetic field. For the determination of the Hall constant the averaged readings for each direction were added so that essentially the effect of the remanence of the magnet (without any current flowing through its winding) was eliminated.

For the measurement of the electrical resistance of each sample in the cold-worked and the annealed states use was made of the tails d and e. They were used instead of the electrodes b and c as the "voltage electrodes" in a Kelvin Bridge arrangement, since the wider spacing between the former increases the accuracy of the measurement. The resistance measurements were made with a current of approximately 50–150 amperes existing between electrodes A and B. A Weston 50-millivolt shunt, "Type 8," set at the 150 ampere position was used as a standard resistor. The rest of the bridge was made up by two Shallcross "decade potentiometers."

The material used for this investigation was commercial electrolytic copper. A chemical analysis[9] made at this laboratory gave the following results: 99.9 percent copper plus undetermined amounts of silver and oxygen. The proportion of ferromagnetic materials was less than 0.007 percent.

Previous to the experiment the sample was annealed for one hour in a helium atmosphere at 595°C. To prevent oxidation, the commercial helium employed was purified by passing it through a liquid nitrogen trap filled with activated charcoal and then over heated copper strips. The samples were then cold-worked in an electrically driven rolling mill.[10]

TABLE I. Hall emf's V_c' and V_c'' obtained by exciting the magnet by currents of opposite direction.

V_c' in microvolts	V_c'' in microvolts
	−56.6
56.0	
	−57.0
55.9	
	−56.7
56.4	
	−56.7
56.4	
	−56.6
56.5	
	−56.5
56.8	
	−56.2
56.5	
	−56.3
56.3	
	−56.4

[9] We are indebted to Mr. D. Ottersen for the analysis.

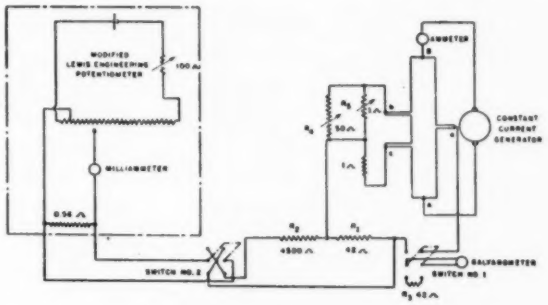


FIG. 2. Electrical circuit used in measuring the Hall effect.

The reduction in cross-sectional area achieved varied between 50 to 75 percent in the preliminary experiments, while the final experiments were performed with 75 percent reduction.

In order to obtain a uniform current density through the sample, the terminals A and B for the primary current were made of heavy copper bars extending over the width of the strip. During the preliminary experiments the bars were soft soldered to the strip. This necessitated taking them off to anneal the strip after the first measurement of the Hall effect, and resoldering them before the measurements of the annealed sample. In order to avoid this complication, the bars were in the final experiments hard soldered to the strips. This caused annealing in parts close to electrodes A and B, but since the strips were approximately 25 centimeters long, and furthermore, cooled by moistened asbestos during the soldering, the annealing was limited to the regions outside the magnetic field. Provisions were made to reproduce the exact position of a strip in the magnet gap before and after annealing.

RESULTS OF PRELIMINARY EXPERIMENTS

In the course of these experiments ten samples were used. Only a very small magnet which produced a field of 4000 gauss over an area of 6 cm by 6 cm was available at the time these experiments were performed. The primary Hall current was only 100 amperes. An average ratio of the Hall constants R_c and R_a of the cold-worked and annealed sample of 1.004 was found. The fluctuations of this ratio from sample to sample were such that not much confidence can be placed in this result. It is stated here in order to support the final contention.

FINAL EXPERIMENTS

All the final experiments were performed under similar conditions using a magnetic field of 9600 gauss and a primary Hall current of 200 amperes. Eleven

[10] We are indebted to Mr. Don Vining for the rolling operation.

TABLE II. Ratios of the Hall coefficients for the cold-worked and annealed states.

1.00186
1.01194
1.00377
1.00172
1.01704
1.01536
1.01577
1.01903
1.0265
0.98801
1.01155

samples were investigated. Table I shows as an example the results of the measurements of the voltage V_c' and V_c'' of one of the samples in the cold-worked state with opposite directions of the magnetic field. Other runs showed similar variations between the individual measurements of V_c and V_a, the voltage measured with the annealed sample.

It can be seen that the individual emf's V_c' and V_c'' obtained for opposite directions of the magnetic field are fairly reproducible. The probable error in the expressions $(V_c)_H=\frac{1}{2}[(V_c')_H-(V_c'')_H]$ and $(V_a)_H=\frac{1}{2}[(V_a')_H-(V_a'')_H]$ is in each case of the order of 0.1 percent. The error associated with the ratio $R_c/R_a=(V_c)_H/(V_a)_H$ as a result of errors in the measurement of the V's is of the order of $\sqrt{2}\times 0.1$ percent ≈ 0.15 percent.

Table II shows the ratio of the Hall coefficients R_c and R_a for the eleven samples constituting the final experiments. The average value of R_c/R_a is 1.010 with a probable error of approximately 0.002 based on the fluctuations of the eleven individual ratios. Only one of these ratios is smaller than one, indicating that in this case the Hall effect increased after annealing. It is also evident that the error in these ratios is considerably larger than the error of 0.15 percent resulting from the fluctuations of the Hall emf's in the course of a single run. This means that errors in this investigation result from either a possible deformation of the strips during annealing or from the uncertainty of repositioning a strip between the magnet poles after annealing, etc.

Efforts were made to find systematic errors which might arise from the annealing. No change in the dimensions could be detected. An inspection of the strips after annealing showed in some cases signs of oxidation. However, the thickness of this layer is negligible, compared to the thickness of the strip and the increased current density in the copper after annealing would have resulted in an increase of the Hall voltage rather than in the observed decrease. The measured change in resistivity of the samples (the average being 2.2 percent) agrees well with the values quoted in the literature. This indicates that our specially shaped samples (thin strips) behaved during the annealing in a way similar to the more bulky ones usually used for resistance measurements.

It is obvious that the change in the Hall constant is so small compared to the probable error that it is rather hopeless to investigate the dependence of the Hall effect on the degree of cold-work.

In the case of eight of the strips, the dimensions were measured in order to obtain the Hall constant for annealed copper. The thickness was measured in several positions across the strip at the position of the electrode. According to these measurements it was found to be $R_a=-5.33\times 10^{-13}$ volt-cm/ampere-gauss compared to previously reported values ranging from minus 4.28 to minus 5.47,[11] the average being minus 5.2×10^{-13} volt-cm/ampere-gauss.

On the basis of the above-quoted theory this change of the Hall effect indicates a decrease of the order of one percent in the number of free electrons as a result of cold-work. If this result is combined with the result of the resistivity measurements, it follows from Eq. (1) that the mean free path is decreased by one to two percent. This decrease was to be expected because of the diminution of the periodicity of the lattice. On the other hand, no such simple explanation is available for the decrease in the number of free electrons. However, the theory refers to the case of a homogeneous metal and even in this case is somewhat over-simplified. In a more refined theory taking into account dislocations, etc., the increase of the Hall effect with the amount of cold-work might be explained in different terms.

We are indebted to Dr. Philip Schwed of this laboratory for valuable discussions.

[11] International Critical Tables (1929), Vol. 6, p. 416.

Structure of Oxide Replicas for Electron Microscopy

SHIGETO YAMAGUCHI
Scientific Research Institute, Komagome, Bunkyo-ku, Tokyo, Japan
(Received December 31, 1951)

The structure of oxide films, which could be used as replicas for electron microscopy, was investigated here. In order to obtain a replica of the metal surface using the oxide film formed on it and isolated from it, the oxide crystals composing the film must be oriented relative to the orientation of the substrate face. The oxide crystals formed on metal, whose atom distances coincided with those of the substrate crystal, could compose a replicating film of the specimen surface. This oxide film did not curl nor expand upon removal from the substrate.

INTRODUCTION

It was reported in a Letter[1] to the Editor by the author, that the oxide replica formed on an iron single crystal was composed of oriented Fe_3O_4 crystallites. It was found in the present study that this orientation did not always take place in the case of NiO replica formed on a Permalloy single crystal. It was established that an oxide replica had to satisfy certain conditions illustrated in this paper in order to give rise to a relieved image in electron microscopy.

EXPERIMENTAL

The finished surface of a single crystal of Permalloy (Ni: 81.5 percent) was etched with an ethanol-bromine solution (10:1 by volume) for about 30 seconds. This etched surface was observed by electron diffraction (Fig. 1). The incident beam in Fig. 1 was parallel to the (011) and (31$\bar{1}$) planes of the face-centered cubic crystal (lattice constant: 3.52A). The macroscopic surface corresponding to the shadow edge appearing in Fig. 1 was inclined by 29 degrees to the (111) plane of specimen crystal.

The etched surface was oxidized in a fused salt bath of $NaNO_3-KNO_3$ (1:1 by weight) at about 300 degrees C for about 5 minutes.[2] The oxide film isolated from the substrate was magnified by means of electron microscopy (Fig. 2(A)). Figure 2(A) shows the octahedral etch figures. The angle between one of the (111) faces found in Fig. 2(A) and the replica film plane was roughly equal to the inclination angle (29 degrees) of the (111) plane of the specimen determined by electron diffraction. Therefore, this oxide film seemed to be a reprint of the specimen surface.

An electron beam transmitted perpendicularly to the oxide film of Fig. 2(A) gave a diffraction pattern which verified the coexistence of NiO crystallites of random arrangement and NiO crystals oriented regularly (Fig. 2(B)). The oxide replica film formed on the etched surface of an iron single crystal gave the micrograph of Fig. 3(A) and the diffraction pattern of Fig. 3(B). The experimental procedures carried out in this case were described in the previous paper.[1] It was remarked in Fig. 2(B) and Fig. 3(B) that the oxide film of Fig. 3 formed on the etched surface of iron was composed of Fe_3O_4 crystals oriented more regularly than the NiO crystals composing the film of Fig. 2. So far as the forms of octahedrons are concerned, the figures of Fig. 2(A) are quite similar to those of Fig. 3(B). It is, however, discernible that the image of Fig. 3(A) exhibits more contrast than that of Fig. 2(A). It had been explained in the previous paper that the orientation of replica film led to the separation of the microscopic image.[1]

DISCUSSION

The orientation of the thin oxide film on metal surface was considered to be conditioned by the following circumstances.

The $Fe^{\cdot\cdot}-Fe^{\cdot\cdot}$ distance found in the Fe_3O_4 crystal formed on iron is 8.37A (lattice constant of Fe_3O_4), and the doubled value of the distance (4.045A) between the nearest two atoms in the (111) plane of iron is $4.045\times2=8.90$A. The difference between these two distances is small, i.e., it amounts to 3.5 percent. This

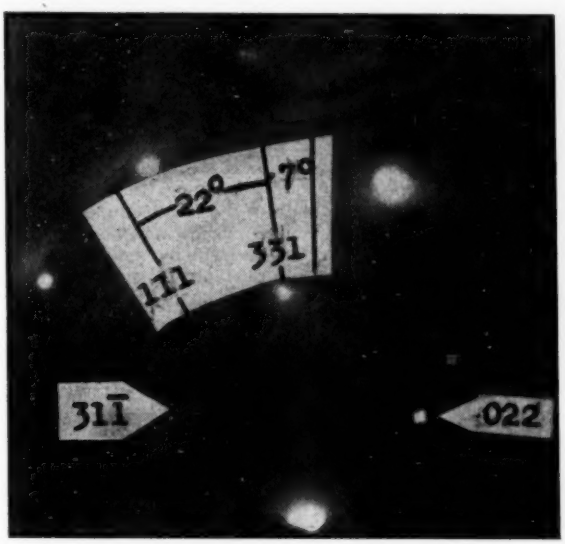


FIG. 1. A diffraction pattern from the etched surface of Permalloy single crystal. The incident beam is parallel to the (011) and (31$\bar{1}$) planes of the face-centered cubic crystal (*i.e.*, along a [$\bar{2}$33] zone axis). The angle between the (111) plane and the macroscopic surface is about 29 degrees.

[1] S. Yamaguchi, J. Appl. Phys. **22**, 1295 (1951).
[2] E. M. Mahla and N. A. Nielsen, J. Appl. Phys. **19**, 378 (1948); S. Yamaguchi, J. Appl. Phys. **22**, 680, 983 (1951).

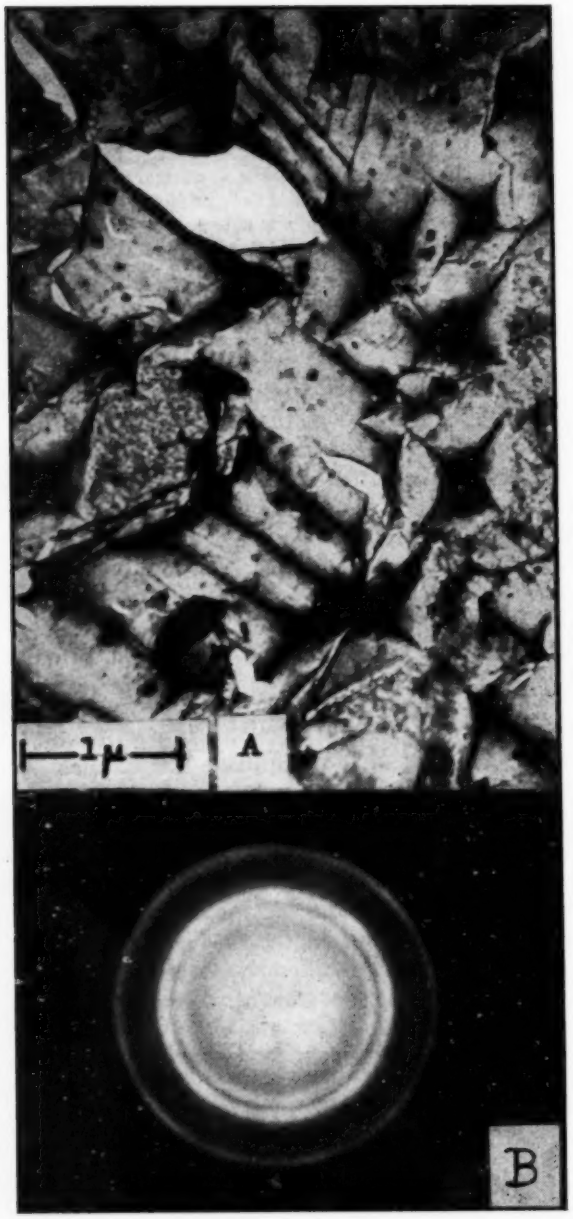


FIG. 2. (A) Micrograph of the oxide film formed on a Permalloy single crystal. The octahedral etch figures are in evidence. (B) Diffraction pattern produced by transmission perpendicular to the film A. This pattern shows the coexistence of the NiO crystals of random arrangement and the crystals oriented regularly.

suggests that the oxide formed on the surface can be naturally oriented according to the orientation of the substrate. As a matter of fact, on the removal of the Fe_3O_4 film from iron, this film was fragile, but it did not curl nor expand.

The $Ni^{\cdot\cdot}-Ni^{\cdot\cdot}$ distance in the (111) plane of NiO is 2.95A, whereas the Ni—Ni distance in the (111) plane of the substrate is 2.49A. The difference between these two distances is considerably larger, i.e., it amounts to 18.5 percent. This means that an oriented oxide layer would be under severe tension on the substrate, and so prefers to form with random orientation. Upon removal of the NiO film from the substrate, the film tended to tear.

Variously oriented facets reproduced in a replica film which is composed of properly oriented crystallites exhibit different refractive indices and, therefore, give an image with higher contrast than the homogeneous film consisting of crystallites oriented at random.

SUMMARY

The oxide replicas prepared on metals and alloys do not always show clearly the topography of the sub-

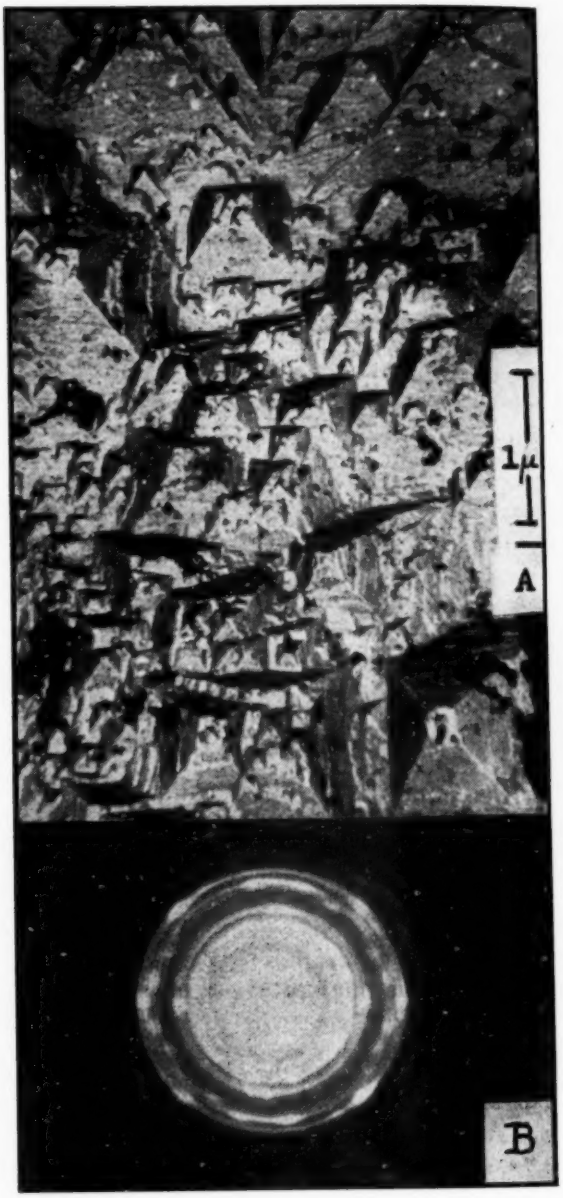


FIG. 3. (A) Micrograph of the oxide film formed on an iron single crystal. This image indicates more relief than that of Fig. 2(A). (B) Diffraction pattern produced by transmission perpendicular to the film A. This pattern shows the Fe_3O_4 crystals oriented regularly.

strata. The Fe_3O_4 formed on iron gave an excellent relief of the surface. The relief indicated in electron images of the NiO formed on Permalloy was comparable with that of Formvar replica. It was concluded that the oxide films formed on metals, whose atom distances coincided with those of the substrata,[3] could reprint the substrate faces satisfactorily. The successful oxide replicas did not expand or curl when removed from the substrate.

ACKNOWLEDGMENT

The Permalloy single crystal used in the present study was prepared by Dr. R. M. Bozorth and his coworkers[4] and given to the author with permission to publish the results obtained. The author wishes to thank Dr. Bozorth for his aid.

[3] G. P. Thomson and W. C. Cochrane, *Theory and Practice of Electron Diffraction* (Macmillan and Company, Ltd., St. Martin's Street, London, England, 1939), p. 181.

[4] Walker, Williams and Bozorth, Rev. Sci. Instr. **20**, 680 (1949).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 4 APRIL, 1952

Flow over Thick Plate with Circular Hole*

W. R. SMYTHE
California Institute of Technology, Pasadena, California
(Received November 8, 1951)

Formulas are derived for the scalar potential of a uniform magnetic field parallel to the plane face of an infinite nonpermeable medium in which a circular hole of radius, a, is drilled normally. Outside the hole the formula is an oblate spheroidal harmonic series for which eleven coefficients are given, and inside it is a Bessel function series for which twenty coefficients are given. The case of a flat-bottomed hole whose depth is greater than its diameter is included. All series converge slowly in the plane of the opening but rapidly elsewhere. When a plane electromagnetic wave falls on such a hole, the fraction of the energy scattered is approximately $11.7(2a/\lambda)^4\cos\theta$ if the wavelength λ is much greater than $2a$ and θ is the angle of incidence. Such a hole in the thick plane wall of a resonant cavity, at a point where H is uniform and E is zero, decreases the resonance frequency ω by $0.2366\mu a^3H^2W_m^{-1}\omega$ where W_m is the magnetic energy in the cavity when H is measured. A lug of radius a affixed to the plane face of a cylindrical rod of the same material, length l and cross section A changes the resistance in the ratio $[1+0.3154a^3/(Al)]^{-1}$.

IN an experimental problem under way in this laboratory it is necessary to know accurately the perturbation of the uniform magnetic-field parallel to the plane thick wall of a resonant cavity when a small circular hole is drilled in it. This is also the problem of an ideal fluid flowing over the infinite plane flange of a circular pipe closed at its far end. Inside the pipe the cylindrical coordinate system ρ, ϕ, z is used and outside the oblate spheroidal system ξ, ζ, ϕ as shown in Fig. 1. The plane of the opening is $z=0$ or $\zeta=0$, the plane outside the opening is $\xi=0$, and the wall of the pipe is $\rho=a$, $z<0$. The relation between the coordinates is

$$z=az'=a\xi\zeta \tag{1}$$

$$\rho=a\rho'=a(1+\zeta^2)^{\frac{1}{2}}(1-\xi^2)^{\frac{1}{2}}. \tag{2}$$

The magnetic scalar potential Ω, whose negative gradient is H, must satisfy Laplace's equation. Suitable solutions for the regions inside and outside the pipe are,[1] respectively,

$$\Omega_i=aH\cos\phi\sum_{n=1}^{\infty}B_n\exp(k_nz')J_1(k_n\rho') \tag{3}$$

$$\Omega_0=aH\cos\phi[-\rho'+\sum_{m=0}^{\infty}A_mQ_{2m+1}^1(j\zeta)P_{2m+1}^1(\xi)]. \tag{4}$$

Equation (4) also holds inside the hole where ζ is negative provided $Q_{2m+1}^1(-j|\zeta|)$ is defined as $-j\pi P_{2m+1}^1(j|\zeta|)-Q_{2m+1}^1(j|\zeta|)$. If $A_0=\pi^{-1}$ and $A_n=0$ when $n\neq0$, then Eq. (4) is the potential when an infinite thin nonpermeable sheet with a hole of radius a in it bounds the field H. The boundary conditions are

$$\partial\Omega_i/\partial\rho=0 \quad \text{at} \quad \rho=a \tag{5}$$

$$\partial\Omega_0/\partial\xi=0 \quad \text{at} \quad \xi=0 \tag{6}$$

$$\Omega_0=-Hx=-H\rho\cos\phi \quad \text{at} \quad \zeta=\infty \tag{7}$$

$$\partial\Omega_i/\partial z=\partial\Omega_0/\partial z, \quad \Omega_i=\Omega_0 \quad \text{at} \quad z=\zeta=0, \quad \rho<a. \tag{8}$$

To satisfy (5) we choose k_n so that

$$J_1'(k_n)=0 \tag{9}$$

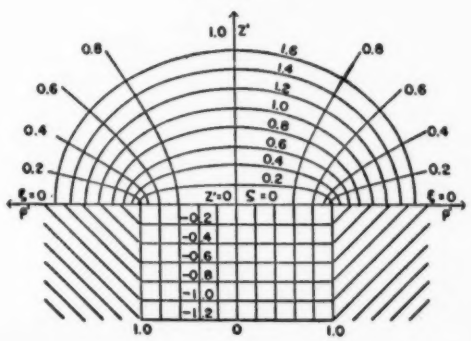


FIG. 1.

* This work was done under ONR contract N6onr24419.

[1] W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1950), second edition, p. 170 and p. 160.

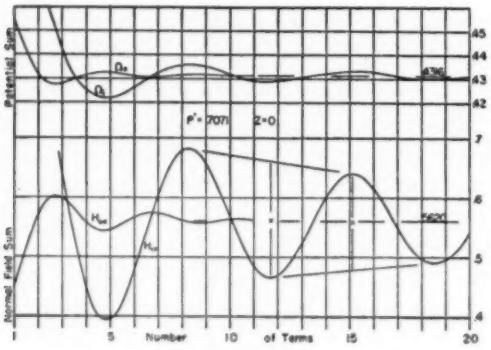


FIG. 2.

The choice of odd order Legendre polynomials[2] in Eq. (4) satisfies Eq. (6) and the vanishing[3] of the summation in Eq. (4) at $\zeta=\infty$ satisfies Eq. (7). A suitable choice of A_m and B_n will satisfy Eq. (8). From the two-dimensional flow around a right angle corner it is known that at a small distance δ from the edge of the hole H varies as $\delta^{-\frac{1}{2}}$ so that the gradients of both Ω_0 and Ω_i are infinite there. The scale factors h_1 and h_2 go to zero as $\delta^{-\frac{1}{2}}$ so that the operators $\partial/(h_1\partial\xi)$ and $\partial/(h_2\partial\zeta)$ in the gradient of Eq. (4) go to infinity as $\delta^{-\frac{1}{2}}$, and the series itself converges there.

EVALUATION OF COEFFICIENTS

The first relation between A_m and B_n is obtained by equating Ω_0 to Ω_i, setting $z=\zeta=0$, multiplying through by

$$\rho' J_1(k_n\rho')d\rho' = \sin\theta J_1(k_n\sin\theta)\cos\theta d\theta, \tag{10}$$

and integrating from $\rho'=0$ to $\rho'=1$ or from $\theta=0$ to $\theta=\frac{1}{2}\pi$. On the left side[4] all terms but the nth vanish. On the right side we have integrals in the form

$$\int_0^{\frac{1}{2}\pi} \sin\theta J_1(k_n\sin\theta)P_{2m+1}{}^1(\cos\theta)\cos\theta d\theta. \tag{11}$$

The Legendre function is written as a polynomial[5] in $\cos\theta$ which yields a set of integrals of the form,[6]

$$\int_0^{\frac{1}{2}\pi} J_\mu(k_n\sin\theta)\sin^{\mu+1}\theta\cos^{2\nu+1}\theta d\theta = 2^\nu\Gamma(\nu+1)k_n{}^{-\nu-1}J_{\nu+\mu+1}(k_n), \tag{12}$$

where $\mu=1$. Thus we obtain

$$B_n' = -J_2(k_n)+\sum_{m=0}^{\infty} A_m Q_{2m+1}{}^1(j\cdot 0) \times\sum_{s=0}^{m}\frac{(-1)^s(4m-2s+1)!!}{(2s)!!(2m-2s-1)!!}\frac{J_{m-s-2}(k_n)}{k_n{}^{m-s}}, \tag{13}$$

$$B_n = 2k_n[(k_n{}^2-1)J_1{}^2(k_n)]^{-1}B_n'. \tag{14}$$

[2] Reference 1, p. 153.
[3] Reference 1, p. 161.
[4] Reference 1, p. 175.
[5] Magnus and Oberhettinger, *Special Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1949), p. 54.
[6] Reference 5, p. 30.

A second relation between A_m and B_n is found by equating $\partial\Omega_i/\partial z$ to $\partial\Omega_0/\partial z$ at $z=\zeta=0$ where ξ is normal to z so that $\partial/\partial z$ may be replaced by $\partial/(\xi\partial\zeta)$ or $a^{-1}(1-\rho'^2)^{-\frac{1}{2}}\partial/\partial\zeta$. From the recurrence formulas[7]

$$[\partial Q_{2m+1}{}^1(j\zeta)/\partial\zeta]_{\zeta=0} = (2m+1)(2m+2)Q_{2m+1}(j\cdot 0). \tag{15}$$

When multiplied through by Eq. (10) and integrated as before, the left side is the same except for the factor k_n, and the right side contains an integral like Eq. (11) but without the factor $\cos\theta$ which was canceled by the $(1-\rho'^2)^{-\frac{1}{2}}$ term. Thus, when the Legendre function is written as a polynomial, (12) appears with even powers of $\cos\theta$ which requires that ν take odd half-integer values and leads to spherical Bessel functions.[8] Thus we obtain

$$B_n' = \sum_{m=0}^{\infty}(2m+1)(2m+2)A_m Q_{2m+1}(j\cdot 0) \times\sum_{s=0}^{m}\frac{(-1)^s(4m-2s+1)!!}{(2s)!!(2m-2s)!!}\frac{j_{m-s+1}(k_n)}{k_n{}^{m-s}}. \tag{16}$$

Elimination of B_n' from Eq. (13) and Eq. (16) gives for each value of n a different linear relation between all the A's so that to obtain A_m exactly requires the solution of an infinite determinant. Actually, the same result can be obtained by another method. The 1×1 determinant is solved for A_0, the 2×2 one for A_0 and A_1, the 3×3 one for A_0, A_1, and A_2, etc. In this way a series of values of A_m is obtained, and it is observed in each case that the successive values, as more terms are taken, approach a definite limit, which is the value one would get if an infinite number were taken. In order to get the same limit to a given number of significant figures, regardless of the extrapolation formula used, it was necessary to calculate the coefficients of the A's to two more significant figures and to verify the solution by substitution into each of the simultaneous equations to the same precision. When this was done it was found that any two-constant extrapolation formula which fitted the last three even terms or the last three odd terms to 5 in the seventh place gave the same limit in

TABLE I.

m	A_m	n	k_n	B_n	n	k_n	B_n
0	+0.22590	1	1.8412	−0.8999	11	33.74	−0.0170
1	−0.0²645	2	5.3314	+0.1645	12	36.89	+0.0151
2	−0.0²143	3	8.5363	−0.0898	13	40.04	−0.0136
3	−0.0³481	4	11.706	+0.0606	14	43.18	+0.0125
4	−0.0³210	5	14.863	−0.0452	15	46.32	−0.0114
5	−0.0³106	6	18.016	+0.0358	16	49.46	+0.0106
6	−0.0⁴578	7	21.164	−0.0295	17	52.61	−0.0099
7	−0.0⁴333	8	24.311	+0.0250	18	55.75	+0.0091
8	−0.0⁴192	9	27.456	−0.0215	19	58.89	−0.0086
9	−0.0⁴114	10	30.602	+0.0190	20	62.03	+0.0081
10	−0.0⁵68		...	...	...	...	...

[7] Reference 1, p. 152, Eq. (6).
[8] Reference 1, p. 186, Eq. (10).

the fifth place. One notes at once that by solving p equations one gets p values of A_0, $p-1$ of A_1, $p-2$ of A_2, etc., so the base for extrapolation of A_m gets shorter as m increases. However, if one writes out the array of values of A_m with p vertical and m horizontal, one finds it easy to extend it uniquely by two rows in the p direction when $m>2$ by diagonal extrapolation of the ratios $A_m^{(p)}/A_{m+1}^{(p+1)}$ which have very regular first and second difference relations.

The same method of solving 1, 2, 3 etc. simultaneous equations was used in calculating the B_n's. In this case even and odd terms lie on the same extrapolation curve and the formula

$$B_n^{(p)}=B_n^{(\infty)}+ap^{-1}+bp^{-2} \tag{17}$$

often fitted the last four or five values of $B_n^{(p)}$ exactly in the seventh place. This formula was also used on all the A_m's although several quite different ones were tried with the same result.

The evaluation of the potential at $z=0$ to the precision with which A_m and B_n are known requires more than ten terms of the spheroidal harmonic series for Ω_0 and more than twenty terms for the Bessel function series for Ω_i. The ratio A_m/A_{m+1} is a smooth curve which is rapidly approaching a constant value at $m=8$ so that A_8, A_9, and A_{10} can be written down with little uncertainty in the last place. A formula for B_n, when $n>8$, was the unanticipated byproduct of an attempt to use the knowledge that when $z=0$, H_z approaches infinity as $(1-\rho'^2)^{-1/3}$ as $\rho'\to 1$. Thus the expression $a\rho'(1-\rho'^2)^{-1/3}$ has the correct behavior at both $\rho'=0$ and $\rho'=1$, and the addition of two small correction terms of the form $b\rho'(1-\rho'^2)^{2/3}$ or $b\rho'$ and $c\rho'(1-\rho'^2)^{5/3}$ or $c\rho'^3$ gives an H_z at $z=0$ that agrees with the spheroidal harmonic series values to the accuracy to which its sums can be found. These forms for the correction terms were used because all can be expanded as a series in $J_1(k_n\rho')$ by the use of Eq. (10), Eq. (11), and Eq. (12). The values of B_n, thus obtained, are a few percent above or below those already found, according to the form chosen for the correction terms, but depend only on a for large values of n. It was noted that when $8\geqslant n\geqslant 4$, the ratio of these values of B_n to those given by Eq. (16) is accurately constant. The assumption that this is also true when $n>8$ yields a formula for B_n'.

$$B_n'\approx k_n^{-7/6}[(0.6261-0.911k_n^{-2}+3.67k_n^{-4})\times\cos(k_n-0.2618)-(0.971k_n^{-1}-1.94k_n^{-3}+4.06k_n^{-5})\sin(k_n-0.2618)]. \tag{18}$$

TABLE II.

ρ'	Potential S.H.	×10⁻⁴	B.F.	×10⁻⁴	Normal field S.H.	×10⁻⁴	B.F.	×10⁻⁴
0.2588	0.15346	±0.3	0.1534	± 2	0.1575	± 3	0.1580	± 10
0.5000	0.29941	±0.4	0.2995	± 2	0.3358	± 2	0.3360	± 10
0.7071	0.4316	±1	0.4317	± 3	0.5620	± 3	0.5620	± 15
0.7880	0.4869	±2	0.4868	± 4	0.7025	±10	0.7020	± 40
0.8660	0.5447	±2	0.5450	±10	0.9080	±20	0.9100	±200

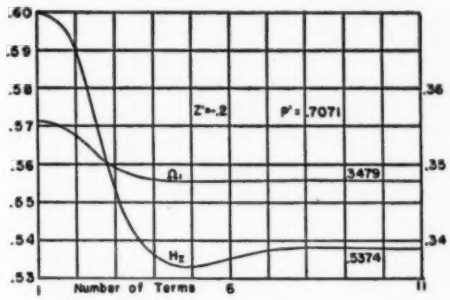


FIG. 3.

The trigonometric terms arise from the use of the asymptotic expansion[9] for the Bessel functions of order $n/3$ that arise from the application of Eq. (12) to terms like $(1-\rho'^2)^{-1/3}$. This result was used for the calculation of B_n when $n>8$. The values appear in Table I. The last digit given for B_n is believed to be significant because when it was dropped and the result rounded off, there was a perceptible mismatch in H_z at some of the check points of Table II.

The foregoing calculations were checked by plotting accurately the values of Ω_0 and Ω_i at $z=0$ given by using p terms in the series of Eq. (3) and Eq. (4) against p for five values of ρ'. For $\rho'=0.7071$ this yields the upper curves in Fig. 2. A similar plot of the z component of H_0 and H_i gives the lower curves. The true value about which each curve oscillates can be estimated in various ways. One used on the Bessel function curves is to join two adjacent crests with a straight line and connect it with a vertical line to the intervening trough. The midpoint of this line lies above the true value because the envelope of the damped wave is convex on the inner side. A similar procedure with two troughs and a crest gives a lower limit. Limits of error for the five points are given after each estimated value in Table II. It will be noted that the boundary conditions of Eq. (8) are satisfied well within these limits so there are no gross errors. Actually the coefficients given for the first eight terms of each series are believed to be correct to one in the last place given.

The rapidity of convergence of the series in Eq. (3) and Eq. (4) and their derivatives increases rapidly with increasing distance from the boundary $z=0$. This is shown in Fig. 3, where at $z'=-0.2$, $\rho'=0.7071$, Ω_i and H_z are plotted as a function of the number of terms taken. For accuracies of 1 percent and 0.1 percent, Ω_i requires one and four terms, respectively, and H_z requires six and eight terms. This point is only one-tenth of a hole diameter inside the opening. This should be compared with Fig. 2.

NUMERICAL ACCURACY CHECK

There is a theorem in current flow which states that[10] if current is flowing in a resisting medium between

[9] Reference 5, p. 22.
[10] Reference 1, p. 233.

fixed perfectly conducting electrodes and the boundary of the medium is modified so as to increase its volume, then the heat generated is increased, decreased, or unchanged accordingly as the potential V, current I, or the product IV is kept constant. In the present case the hole adds to the volume occupied by the magnetic field. If the distant field and potential are kept constant, the net change in the field energy should be zero so that

$$\int_{v_0}(\nabla\Omega_0\cdot\nabla\Omega_0-H^2)dv+\int_{v_i}\nabla\Omega_i\cdot\nabla\Omega_i dv=0, \quad (19)$$

where Ω_0 and Ω_i are given by Eq. (4) and Eq. (3). Both terms involve the integration of the product of two series. Because of the orthogonal properties of Legendre functions, Bessel functions, and the associated surface vector functions,[11] the crossproduct terms are eliminated by the ξ-integration in one term and by the ρ'-integration in the other. The result is, after constant factors are canceled out,

$$-\frac{4A_0}{3}+\sum_{m=0}^{\infty}\frac{(2m+1)^2(2m+2)^2\pi}{2(4m+3)}A_m^2+\sum_{n=1}^{\infty}\frac{k_n^2-1}{2k_n}[J_1(k_n)]^2B_n^2=0; \quad (20)$$

when the numerical values of Table I are inserted, this becomes

$$-0.3012+0.1086+0.1924=-0.0002\approx 0. \quad (21)$$

This is a stringent test since all coefficients are squared, thus doubling the errors, and all terms are positive so that systematic errors add. Forty terms were used in the Bessel function series which converges slowly. When $n>20$ only the first term in Eq. (18) is needed; the cosine is constant because successive values of k_n differ by π and $k_n^{-7/6}$ can be found with a slide rule. The numbers corresponding to Eq. (21) for the thin sheet are $(3\pi)^{-1}(-4+2+2)$.

FIELD ON AXIS

For some purposes knowing the field on the axis is important. Inside the hole at a distance z from the opening this is found from Eq. (3) by finding $-(\rho' a)^{-1}\partial\Omega_i/\partial\phi$, setting $\phi=\frac{1}{2}\pi$ and letting $\rho'=0$. Thus

$$|H_z|_{axis}=\tfrac{1}{2}H\sum_{n=1}^{\infty}B_n k_n e^{-k_n|z|/a}, \quad (22)$$

where k_n and B_n are given in Table I. This converges very rapidly if z is not too small. The field on the axis outside the hole can be found by expressing Eq. (4) in cylindrical coordinates. Starting with a formula in Hobson[12] one can show that

$$Q_n^1(j\zeta)P_n^1(\xi)=\frac{n(n+1)}{2}\times\sum_{p=0}^{\infty}\frac{(-1)^p\rho'^{2p+1}}{2^{2p}p!(p+1)!}\frac{d^{2p+1}Q_n(jz')}{dz'^{2p+1}}. \quad (23)$$

When $\rho'=0$ only the first term survives, and when this is substituted in $-(\rho' a)^{-1}\partial\Omega_0/\partial\phi$ with $\phi=\frac{1}{2}\pi$ there results

$$|H_z|_{axis}=H[1-\sum_{m=0}^{\infty}A_m(m+1)(2m+1)dQ_{2m+1}(jz')/dz']. \quad (24)$$

This derivative is tabulated[13] for $m=0, 1, 2, 3, 4$. Figure 4 is an accurately drawn curve plotted from Eq. (22) and Eq. (24) showing the transverse magnetic field on the axis near the opening. It will be seen that the field falls to half value a distance $0.149z/a$ inside

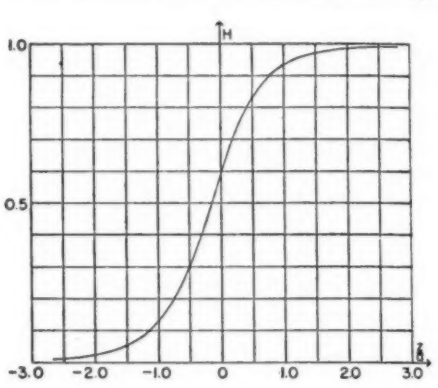


FIG. 4.

the hole. It is almost, but not quite, symmetrical about this point. When the lower half is rotated 180° about the $0.5H$ point, it falls a maximum distance $0.015H$ below the upper half at $0.75z/a$ but coincides again at $2.0z/a$.

FLAT-BOTTOMED HOLE

Suppose that parallel and equal magnetic fields are established on opposite sides of an infinite plane slab of nonpermeable material of thickness $2b$ in which there is a circular hole of radius a normal to both faces. Since in Eq. (22) the transverse field at $z'=-4$ is less than one thousandth that at $z'=0$, it is evident that when $2b\geqslant 4a$ the field outside the hole is still given to 1 part in 1000 by Eq. (4). From symmetry the resultant field in the hole is tangential to a plane parallel to and equidistant from the faces of the slab. This is

[11] Reference 1, p. 151 and p. 176.

[12] E. W. Hobson, *Spherical and Spheroidal Harmonics* (Cambridge University Press, London, England, 1931), p. 423 (bottom of page).

[13] *Mathematical Tables Project, Tables of Associated Legendre Functions* (Columbia University Press, New York, 1945).

exactly the boundary condition for a hole of depth b with a flat bottom. Thus if we let $z'=(z''-b)/a$ so that z'' is measured from the bottom of the hole and superimpose two solutions like Eq. (3), we get

$$\Omega_h=2aH\cos\phi\sum_{n=1}^{\infty}B_n e^{-k_n b/a}\cosh(k_n z''/a)J_1(k_n\rho/a). \quad (25)$$

This is the potential inside a hole whose depth b is greater than $2a$.

SCATTERING FROM CIRCULAR HOLE IN MIRROR

The propagation equation of any of the field vectors in an electromagnetic wave is

$$\nabla^2 B=-\omega^2\mu\epsilon B=-4\pi^2\lambda^{-2}B. \quad (26)$$

If the wavelength is long compared with the dimensions of the region involved, the right side approaches zero and Eq. (26) becomes Laplace's equation. Thus if a plane wave falls normally on the infinite flange of a circular pipe, all surfaces being perfectly conducting, the values of the magnetic fields obtained by differentiation of Eq. (3) and Eq. (4), multiplied by $\cos\omega t$ will be correct near the opening. From these the complete multipole radiation scattered by the hole can be calculated. Only the most important dipole term will be worked out here. If the series in Eq. (24) is expanded in reciprocal powers of z, the first term[14] is

$$H_1=0.1506Hz'^{-3}=0.1506H(a/z)^3. \quad (27)$$

The field in and normal to the equatorial plane close to a magnetic loop dipole of radius b carrying a current I is[15]

$$H=\tfrac{1}{4}Ib^2r^{-3}\cos\omega t. \quad (28)$$

These are equivalent if $\frac{1}{4}Ib^2=0.1506Ha^3$. Substitution of this value in the expression for the distant field[15] gives

$$E_\phi=(\mu\epsilon)^{-\frac{1}{2}}B_\theta=-0.3012\omega\beta B_0 a^3 r^{-1}\times\sin\theta'\cos(\omega t-\beta r), \quad (29)$$

where B_0 is the amplitude of the magnetic induction in the incident traveling wave. This is half that of the standing wave magnetic induction at the surface of the plate. The angle, θ', is measured from the loop axis which parallels the original magnetic field. The ratio of the integral of Poynting's vector $\frac{1}{2}E_\phi H_\theta$ over a large hemisphere to the integral, $\frac{1}{2}\pi a^2 E_y B_x$, of the incident Poynting vector over the hole is

$$R=11.7(d/\lambda)^4, \quad (30)$$

where d, the diameter of the hole, is much less than λ. If the angle between the incident ray and the axis of the hole is θ and if the magnetic vector lies in the plane of incidence, then the B_0 in Eq. (29) must be multiplied by $\cos\theta$ to get the tangential B. But the energy striking the hole area is also multiplied by $\cos\theta$ so that the ratio of the amount scattered to that incident is

$$R_\theta=11.7\cos\theta(d/\lambda)^4. \quad (31)$$

If the electric vector lies in the plane of incidence, then there is an electric field normal to the surface so this case cannot be solved here.

EFFECT OF HOLE ON CAVITY RESONANCE FREQUENCY

The theory of resonant cavities shows that if a hole is drilled in the wall, the change $\Delta\omega$ in the resonance frequency is[16]

$$\Delta\omega=-\tfrac{1}{4}jW_m^{-1}\int \mathbf{E}_t\times\mathbf{H}_0\cdot\mathbf{n}dS, \quad (32)$$

where W_m, $\mathbf{E}_t$, and $\mathbf{H}_0$ are the instantaneous phasor values of the magnetic energy in the cavity and the tangential components of the actual $\mathbf{E}$ and of the $\mathbf{H}_0$ with no hole. The integration is over the plane of the opening. In our case $\mathbf{H}_0$ is uniform and in the x direction and $\mathbf{n}$ is in the z direction so the integral of E_y over the hole is required. Integration of the z component of the Maxwell equation that relates the curl of $\mathbf{E}$ to the rate of change of $\mathbf{B}$ gives

$$E_y=j\omega\int B_x dx+\int(\partial E_z/\partial y)dx. \quad (33)$$

To get $B_z=\mu H_z$ we differentiate Eq. (4) with respect to z and set $\zeta=0$. From Eq. (1) $dz=\xi d\zeta$ when $\zeta=0$ and from Eq. (2), when y' is constant,

$$\rho'^2=x'^2+y'^2=1-\xi^2, \quad -\xi d\xi=x'dx'=(1-\xi^2)^{\frac{1}{2}}\cos\phi dx'.$$

Keeping in mind that

$$P_{2m+1}{}^1(\xi)=(1-\xi^2)^{\frac{1}{2}}dP_{2m+1}(\xi)/d\xi,$$

we see that the integrand in (33) becomes

$$[dP_{2m+1}(\xi)/d\xi]d\xi,$$

so that

$$j\omega\int B_x dx=j\omega\mu aH\times\sum_{m=0}^{\infty}A_m[dQ_{2m+1}{}^1(j\zeta)/d\zeta]P_{2m+1}(\xi), \quad (34)$$

where the integration was from the edge of the hole where B_x and $P_{2m+1}(\xi)$ are zero to ξ. From Eq. (2) the area element is

$$dS=a^2\rho'd\rho'd\phi=-a^2\xi d\xi d\phi=dxdy.$$

When Eq. (34) is multiplied by dS and integrated from $\xi=0$ to $\xi=1$ and from $\phi=0$ to $\phi=2\pi$, only the $m=0$ term survives because $\xi=P_1(\xi)$. Since from symmetry

[14] Reference 5, bottom of p. 55.
[15] Reference 1, p. 489.
[16] Reference 1, p. 539.

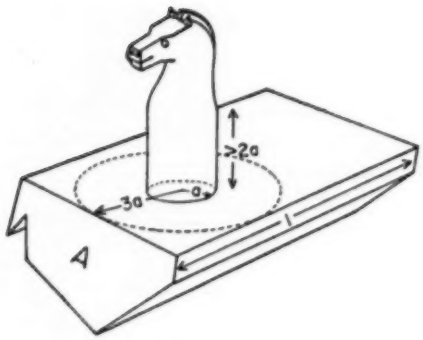


FIG. 5.

$\partial E_x/\partial y$ is an odd function of y, its integration across the hole in the y direction is zero. Thus writing -2 for the ζ-derivative when $\zeta=0$ and 0.2366 for $\frac{1}{3}\pi A_0$ gives

$$\Delta\omega=-0.2366\mu a^3H^2W_m^{-1}. \tag{35}$$

This is the change $\Delta\omega$ in the resonance frequency ω of a cavity when a hole of radius a is drilled in the plane wall whose thickness is greater than $2a$ at a point where H is uniform and E is zero. This implies that a is small compared with the cavity wavelength. The formula should also hold quite well when the wall is curved provided its radius of curvature is much larger than a.

EFFECT OF CYLINDRICAL SIDE LUG ON BAR RESISTANCE

When V_i, V_0, and τI are written for Ω_i, Ω_0, and μH in Eq. (3) and Eq. (4), they represent the potential when current flows in a medium of resistivity τ with a plane boundary to which is affixed normally a solid cylinder of resistivity τ. The current density $\mathbf{i}$ is uniform and parallel to the surface when ζ is large. To find the effect of the side lug on the resistance, one compares the power dissipation inside the hemispheroid ζ_0 with and without the lug, keeping $\mathbf{i}$ constant outside it. Setting $\Psi=\Phi=V$ in Green's theorem[17] gives the power dissipation to be

$$\tau^{-1}\int V(\partial V/\partial n)dS=\tau^{-1}\int(\nabla V)^2dv=P. \tag{36}$$

At a great distance from the hole all but the $m=0$ term in Eq. (4) are negligible, and the spheroid becomes spherical[18] so that Eq. (4) becomes

$$V_0=\tau i(-r+\tfrac{2}{3}a^3r^{-2}A_0)\cos\phi\sin\theta. \tag{37}$$

Substitution in Eq. (36) and integration over the range $\theta=0$ to $\theta=\frac{1}{2}\pi$ and $\phi=0$ to $\phi=2\pi$ give, since $\partial/\partial n=\partial/\partial r$ and $dS=r^2\sin\theta d\phi d\theta$,

$$P_L=\tfrac{1}{4}\tau i^2\pi^2r^2(r+\tfrac{2}{3}a^3r^{-2}A_0). \tag{38}$$

The ratio of the power with the lug to that without ($A_0=0$) is

$$P_L/P_0=1+\tfrac{2}{3}a^3r^{-3}A_0=1+4\pi a^3A_0/(9v), \tag{39}$$

where v is the volume of the hemisphere. The power dissipation density is constant outside the hemisphere so Eq. (39) holds for any shaped volume v that includes it. In particular, consider a cylinder of constant cross-sectional area A and length l which has a flat side as shown in Fig. 5. Power loss is inversely proportional to resistance when current is constant so that Eq. (39) gives

$$R/R_L=1+0.3154a^3A^{-1}l^{-1}. \tag{40}$$

The tables[13] show that

$$Q_1^1(j\cdot3)[Q_1^1(j\cdot3)]'/\{Q_1^1(j\cdot0)[Q_1^1(j\cdot0)]'\}=0.000933,$$

so that the integrand in Eq. (36) has fallen to 1/1000 of its maximum value when $\zeta=3$. This is approximately a hemisphere of radius $3a$ centered at the origin. Thus, if all the boundaries in Fig. 5 are more than $3a$ distant from the center of the lug junction, Eq. (40) will be quite accurate. The side lug may have any shape or size at a distance $2a$ or more from the junction without affecting the field in the main bar by as much as 1 part in 1000.

[17] Reference 1, p. 53, Eq. (3).

[18] Reference 1, p. 161 and p. 149.

Interferometric Analysis of Airflow about Projectiles in Free Flight

F. D. BENNETT, W. C. CARTER, AND V. E. BERGDOLT
Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland
(Received July 19, 1951)

A description is given of the adaptation to the ENIAC of the interferogram reduction problem for axisymmetric flows. Isopycnal charts from interferograms of projectiles in free flight are presented, and comparison is made with similar studies by other workers using fixed models in a wind tunnel or a circular jet. In general, good agreement is found between density fields obtained by the two methods. Sources of error are discussed; and in particular, it is found that random errors accumulate to effect adversely the density values near the axis of the flow.

I. INTRODUCTION

WHILE interferometric techniques for the study of high speed gas flows have been known for many years, the use of this method on any appreciable scale has had to await the development of high speed calculating machinery. The analysis of an interferogram, i.e., calculation of lines of constant density from fringe shift values, of a two-dimensional flow may be accomplished rather simply; for curves of constant density (isopycnals) are curves of constant fringe shift and setting the interferometer for uniform field will result in an interferogram in which recorded fringes form isopycnal lines. If the disturbance is three-dimensional, the resulting integral equation relating fringe shift to air density in the flow can at present be solved only for the case of axisymmetric flows. As the solution itself involves an integration for every radius in a z-plane perpendicular to the axis of the flow, the numerical approximation of the integral for each radius of a given z-plane and the analysis of a field of flow at several hundred points is a major undertaking.

With the development of high speed calculating machinery such as the IBM machines and ENIAC, the situation has undergone considerable changes.

Within the past three years several papers based on wartime research by Ladenburg and collaborators[1–3] have appeared. Complete reductions of three-dimensional interferograms are given and the study of phenomena in axisymmetric jets carried out without severe limitation because of the difficulty of reducing and interpreting results. These authors make use of hand computation methods at the outset and later of IBM reduction techniques developed by H. Polachek and described by F. J. Weyl.[4]

No comparable studies have yet appeared of flows around axisymmetrical projectiles in free flight, although the method has been applied in restricted fashion to the analysis of the flow around small-caliber projectiles by Tesson[5] who cites earlier work of Cranz* and collaborators.

In this report we wish to describe our recent attempts to apply the interferometer method to the study of air flows about supersonic, axisymmetric projectiles.

II. THEORY OF THE REDUCTION PROCESS

If one considers the formation of fringes by a Mach-Zehnder interferometer oriented with its disturbed beam passing at right angles to the axis of an axisymmetric flow, one is led to an integral equation expressing the dependence of fringe shift δ within the disturbed region upon the density difference $\rho-\rho_0$ within the flow induced by the disturbance.†

$$\delta(u_i, z)=\frac{K}{\lambda^*}\int_{u_i}^{u_N}\frac{(\rho-\rho_0)dv}{(v-u_i)^{\frac{1}{2}}}, \tag{1}$$

where r, z, θ are cylindrical polar coordinates with z the axis of the disturbance, $v=r^2$, $u_i=y^2$, and $u_N=r_N^2$, r_N refers to the outer radius of the disturbance and y to some interior value for which the fringe shift is given by (1) while K is the Gladstone-Dale constant, and λ^* is the wavelength of the light in vacuum.

Although Eq. (1) may be used to obtain density values following a method involving simultaneous linear equations explained by H. Schardin[6] and described in Weyl's paper, a form more useful for our purposes can be obtained by solving the integral equation for $\rho-\rho_0$ using a method by Abel. One obtains,

$$(\rho-\rho_0)_{u_i}=\frac{-\lambda^*}{\pi K}\int_{u_i}^{u_N}\frac{\frac{d\delta}{dv}dv}{(v-u_i)^{\frac{1}{2}}}. \tag{2}$$

[1] Ladenburg, Winckler, and Van Voorhis, Phys. Rev. **73**, 1359 (1948).
[2] Ladenburg, Van Voorhis, and Winckler, Phys. Rev. **76**, 662 (1949).
[3] J. Winckler, Rev. Sci. Instr. **19**, 307 (1948).
[4] F. J. Weyl, Naval Ordinance Report 211-45 (December, 1945).
[5] M. Tesson, Etude *G*, *b*, Note No. 2 (April 15, 1940) Service Technique des Fabrications D'Armement, Section Balistique Exterieure (Paris, France).
* The reference to Cranz and various older works dealing with the interferometer and interferometry of gas flows are given in references 1–4. The reader interested in the background literature should consult these papers.
† For a complete derivation of (1) and its solution see Appendix I or the report by F. J. Weyl.
[6] H. Schardin, Z. Instrumentenk. **53**, 396, 424 (1933).

If one assumes next that for sufficiently small ranges, $u_j \leq u \leq u_{j+1}$,

$$\frac{d\delta}{du} = \frac{\delta_{j+1} - \delta_j}{u_{j+1} - u_j}, \tag{2a}$$

(2) becomes

$$(\rho - \rho_0)_i = \frac{-2\lambda^*}{\pi K} \sum_{j=i}^{N-1} \frac{\delta_{j+1} - \delta_j}{(u_{j+1} - u_i)^{\frac{1}{2}} + (u_j - u_i)^{\frac{1}{2}}}. \tag{3}$$

This is an approximation to the integral (2) which can be easily adapted to numerical evaluation by ENIAC. If one has a table of fringe shift values *versus* radii for some trace $z = \text{const}$, the corresponding values of $(\rho - \rho_0)_i$ may be deduced from (3) subject only to the restriction that the radii defining the various zones be chosen close enough together so that the approximation $d\delta/du = \text{const}$ is well fulfilled. For the present it is sufficient to observe that very good accuracy is obtained with zones separated by about 0.1 mm in a field where r_N, the outer radius, is of the order of 1 cm.

Equation (3) is useful for continuous density distributions but breaks down when one or more step discontinuities in density occur. As the shock waves which bound a supersonic flow or occur within it very closely approximate step discontinuities in density, a much more valuable form is obtained by use of the "reduced" functions which are defined as follows.

Thus,

$$\left.\begin{aligned} (\rho - \rho_0)_{\text{red}} &= (\rho - \rho_0) - \rho_2 \\ \delta_{\text{red}} &= \delta - \rho_2'(r_s^2 - r^2)^{\frac{1}{2}} \end{aligned}\right\} r < r_s, \tag{4}$$

and

$$\left.\begin{aligned} (\rho - \rho_0)_{\text{red}} &= (\rho - \rho_0) \\ \delta_{\text{red}} &= \delta \end{aligned}\right\} r_s \leq r \leq r_N,$$

with $\rho_2' = \rho_2 2K/\lambda^*$ and ρ_2 the step discontinuity in density at radius $r_s \leq r_N$. The theory of this reduction has been given by F. J. Weyl[4] and is developed in Appendix I. With these reduced functions and $u = r^2$, (3) becomes upon division by ρ_0,

$$(\rho/\rho_0)_{\text{red}\,i} = 1 - \frac{2\lambda^*}{\pi K \rho_0} \sum_{j=i}^{N-1} \times \frac{\delta_{j+1} - \delta_j - \rho_2'[(r_s^2 - r_{j+1}^2)^{\frac{1}{2}} - (r_s^2 - r_j^2)^{\frac{1}{2}}]}{(r_{j+1}^2 - r_i^2)^{\frac{1}{2}} + (r_j^2 - r_i^2)^{\frac{1}{2}}}. \tag{5}$$

1. ERROR IN THE NUMERICAL APPROXIMATION

Several expressions are derived in Appendix I for the error E in the approximation to the integral. We quote here,

$$E = \frac{-2\lambda^*}{\pi K \rho_0} \sum_{j=i}^{N-1} \int_{u_j}^{u_{j+1}} \left(\frac{d^2\delta}{du^2}\right)_{u(\theta)} (v - a)\, d(v - u_i)^{\frac{1}{2}}, \tag{6}$$

with a and $u(\theta)$ in the interval (u_j, u_{j+1}). Equation (6) shows that where $d^2\delta/du^2$ is large, and the curvature of $\delta(u)$ is usually large, the largest contributions to the error will occur. This is important for continuous $(\rho - \rho_0)$ distributions but particularly so for density functions with a step discontinuity; for $d^2\delta/du^2$ and $d\delta/du$ both go negatively infinite at the discontinuity. Weyl[4] has shown by a particular example, the startling improvement that can be obtained through use of the reduced functions when a step in $(\rho - \rho_0)$ is present. The reason for this may be seen by inspection of Eqs. (I-7) and (I-9) of Appendix I. $d^2\delta_{\text{red}}/du^2$ suffers no infinity as $u \to u_s$; whereas $d^2\delta/du^2 \to \infty$ as $(u_s - u)^{-\frac{1}{2}}$ in the limit.[7]

The use of the reduced functions then eliminates both the infinity and the error resulting from the step; however another possibility for overshoot exists which has apparently been overlooked until now. From (I-7) one may see that $d^2\delta_{\text{red}}/du^2$ has an infinity at $u = u_N$ if the first derivative of $(\rho - \rho_0)_{\text{red}}$ at $u = u_N$ is different from zero. Since the situation usually encountered thus far is that $u_s = u_N$, a slight overshoot at the shock wave may often be interpreted as faulty choice of the ρ_2' value, but the form of (I-7) indicates clearly that this interpretation does not necessarily follow. With nonzero first derivative of $(\rho - \rho_0)_{\text{red}}$ at $u = u_N$, an overshoot may occur even when the density step has been well accounted for.

In some of the trace reductions made at this laboratory, particularly for traces near the tip of the projectile, a small overshoot is encountered within 2–3 zones of the shock wave even when reduced functions are used. Investigation of $d(\rho - \rho_0)_{\text{red}}/du$ for Taylor-Maccoll flow shows that the largest values occur near the tip where the derivative becomes infinite. For the most part the overshoot can be ignored in fairing in the density curve. Should intensive study of a tip region become desirable or the overshoot become troublesome elsewhere along the shock, it may be removed by defining a new set of reduced functions by adding $A(v - u)$ and $\frac{2}{3}AK/\lambda^*(u_N - u)^{\frac{1}{2}}$ to $(\rho - \rho_0)_{\text{red}}$ and δ_{red}, respectively. A is a constant equal to twice the derivative of $(\rho - \rho_0)_{\text{red}}$ at $u = u_N$. Before leaving the subject of the reduced functions, mention should be made of another important application where their use removes overshooting of the density curve.

Just within lines extending the boundaries of a supersonic, cal 30 projectile rearward beyond the truncated base, lies a turbulent wake region of approximately conical shape with apex toward the rear. Here one would expect an expansion fan resembling Prandtl-Meyer flow. Hence, there should be sharp density gradients on traces near the base. Along a trace just behind the base, the reductions show that the integrated mean density changes very suddenly within a few tenths of a millimeter from values about $0.9\rho_0$ to about $0.3\rho_0$. While this

[7] This discussion assumes the continuity of $(\rho - \rho_0)_{\text{red}}$ and its first two derivatives for $u \leq u_N$.

is not a shock region the change in density is nevertheless violent enough to necessitate use of the reduced functions; otherwise an undershoot occurs which sometimes gives negative values of the density.

For large projectiles with well-developed boundary layer this undershoot may not occur, but for the cal 30 and cal 22 projectiles observed so far the boundary layer is undetectable on the interferogram and serious errors result unless appropriate reduced functions are used.

2. EFFECT OF ZONE SIZE ON THE ERROR

The effect of zone size on error is especially clear in the case of conical flow. A conical flow, approximated in practice by the flow about cone or cone-cylinder projectiles, may be defined by the invariance of density, pressure, and velocity components on radial lines through the vertex of the flow. Such flows have especially simple interferometric properties as has been discussed previously.[8] Thus density is constant under the conical transformation T which carries (x, y, z) into (tx, ty, tz) on a line through the vertex. As a consequence, the bound (6) for error (E) remains invariant under the transformation. This may be easily verified by noting that in (6), δ transforms as t; so the second derivative transforms as $1/t^3$ while the multiplying factors transform as t^3. Now zone width $w_j=r_{j+1}-r_j$ transforms according to $w'=T[w]=tw$; so w' and t approach zero together. This means that for density determinations to be of comparable accuracy on separate traces in a conical flow the zone widths must conform to T. In practice one usually employs a constant minimum zone width for all traces; thus traces near the vertex are much more roughly divided than those far away and the bound (6) varies from trace to trace. As the vertex is approached the determinations become less accurate until finally the methods fails altogether.

III. ADAPTATION OF THE CALCULATION TO ENIAC

The logical basis for coding problems for high speed computing machines has been described elsewhere.[9,10] The calculation problem posed by Eq. (5) offers no novel features in so far as ENIAC is concerned.

The complete program is kept on file in the Computing Laboratory, BRL under the designation ENIAC PROBLEM NO. 49.1. Persons interested in studying the program or in having data reduced by ENIAC should request permission from Director, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland.

With the present ENIAC program several interferograms have been reduced in the past year. While the discussion of these will be deferred to a later paragraph, it is worth noting here that the saving in computation time is very great. In a recent example, complete reduction of two sides of an interferogram (regarding each side of the centerline as a separate problem) using 25 traces 6 of which occurred back of the base, took about 7 shifts of ENIAC time and resulted in density values at some 3200 points. Each of the base traces was divided into approximately 150 zones thus giving rise to a calculation that cannot be performed with existing 60 to 100 zone tables and IBM machines, and certainly would not be feasible for hand computation.

IV. MEASUREMENT OF INTERFEROGRAMS

The technique of obtaining interferograms of projectiles in free flight in the model range has been described before.[11]

For the purpose of measurement two Zeiss, tool maker's microscopes have been available. These have screws calibrated to 1.0 and 0.1 micron, respectively; while measurements of position are recorded to this full sensitivity, no claim may be made that the accuracy of measurement is this good. A single observer can reproduce settings on a fringe edge to within 10μ most of the time, but two observers rarely agree as the location of a fringe boundary to better than 20μ. It is felt that the erratic error of setting is not greater than 10μ for measurements made by a single observer.

Since the interferometer has been adjusted to form fringes parallel to the axis of the range and since at Mach numbers greater than 1 the projectile drops less than 0.06 in. in a flight of 20 ft (thus making an angle with the axis of the range of less than one minute of arc) the assumption is made that a nonyawing projectile will travel along the axis of the range and should therefore be accurately parallel to the fringes. Accordingly, the interferogram is aligned on the microscope stage so that the fringe passing nearest the centerline of the projectile is parallel to the crosshairs and to the z motion of the microscope stage.

In order to determine whether the projectile is yawing, the first procedure is to determine the centerlines of both the projectile and shock wave. This is done by measuring the intersections of the projectile boundary and shock wave with traces $z_1=\text{const}\cdots\text{etc}$. Taking the mean of the upper and lower intersections for traces $z_1, z_2, \cdots z_m$ gives a set of m points defining the axes of the projectile and a similar set of points for the shock wave. Once these axes are known their slopes with respect to the horizontal fringes may be found. This quantity for the projectile axis is the apparent yaw of the projectile. If the base of a cylindrical projectile appears sharp and square on the interferogram the inference may be made that the yaw is almost entirely in the plane of the interferogram. For the most stable rounds fired to date apparent yaws ranging from 20–40

[8] Giese, Bennett, and Bergdolt, J. Appl. Phys. **21**, 1226 (1950).
[9] R. F. Clippinger, BRL Report No. 673 (Aberdeen Proving Ground, Maryland, 1948).
[10] H. H. Goldstine and J. von Neumann, *Planning and Coding of Problems for an Electronic Computing Instrument* (Institute for Advanced Study, Princeton, New Jersey, 1947).
[11] Bergdolt, Greenwald, and Sleator, Phys. Rev. **76**, 879 (1949) (A); BRL Mem. Rep. No. 548 (May, 1951).

minutes of arc have been obtained. Rounds are considered suitable for reduction if the apparent yaw is roughly the actual yaw and not larger than approximately half a degree.

A plot of the position of the shock wave is made and a graphical smoothing of the data employed in order to reduce scatter in the intersection points of traces with the shock wave.

With this preliminary data in hand, measurements of fringe shift data are obtained as follows.

a. Undisturbed fringe number N_u *vs* y is obtained along the trace z_1=const in the undisturbed region outside the shock wave.

b. The edges of disturbed fringes are measured tabulating fringe number N_D *vs* y from the shock wave to the projectile boundary (or axis of projectile if back of the projectile base).

c. The correspondence between N_u and N_D is obtained by following a fringe through the shock wave near the nose of the projectile. This procedure is always possible for axisymmetric disturbances because the fringes are continuous through the shock wave as may be seen from Eq. (9) below; for two-dimensional disturbances the fringes are discontinuous and white light fringes must be used.

d. The process is repeated for $z_2, z_3 \cdots z_m$ traces on both sides of the projectile center line.

e. Measurements along the lengths of undisturbed fringes may be taken in the undisturbed region i.e., y *vs* z with N_u=const.

f. Measurements at corresponding $z_1', z_2' \cdots z_m'$ of N_u *vs* y may be taken from an undisturbed interferogram. Because of the blurring of fiducial marks in the interferogram produced by the rotating mirror, this method of obtaining undisturbed fringe positions has not been used very much. Difficulties are encountered in trying to locate corresponding points in disturbed and undisturbed interferograms.

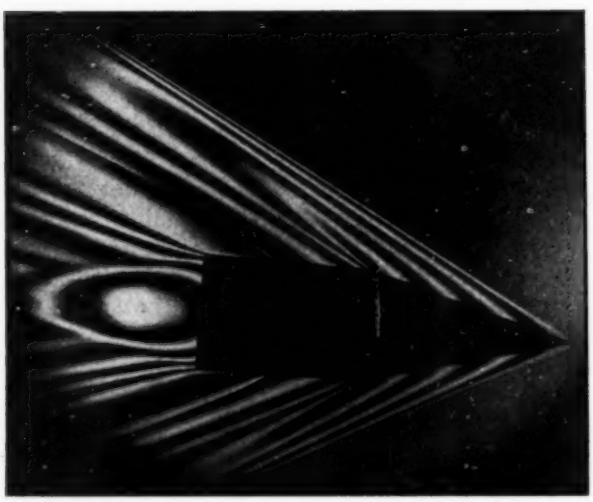


FIG. 1. Single fringe interferogram of 15° half-angle, cal 30 cone cylinder at $M=2.7$, $p=1$ atmos.

g. A 10×–12× enlargement of the interferogram is often useful, as measurements corresponding to those outlined above may be made with a hand rule for purposes of checking.

Several methods of obtaining fringe shift δ from the measurements have been tried of which the most satisfactory is a graphical one.

The procedure is to plot both N_D and N_u *vs* y for z_j=const $j=1, 2, 3, \cdots m$ with y scale large enough so that 1 mm=10μ and 1 mm=0.01 on the N scale. The N_u line is extrapolated into the disturbed region either by extending the line of N_u points found outside the shock wave, or as in *e*, extrapolating points of N_u into the disturbance horizontally along z to help fix the N_u line. Points taken from measurement of traces in an undisturbed interferogram may be used here also. Finally, some success has been had in extrapolating the N_u line along y from the shock wave using both the external measured points and an averaged slope value found from measurements on traces in an undisturbed interferogram as in *f*. Of all methods tried this last is probably the least subject to errors of extrapolation.

Fringe shifts $\delta=N_D-N_u$ for traces z_j=const may then be read off the graph for as many values of y as necessary to obtain reasonable accuracy in the numerical integration. This is always a larger number of points than would be obtained if only the N_D values corresponding to fringe edges were used.

If the fringes are not composed of equal, symmetrical, dark and light bands, and this is usually the case, it is preferable to average successive pairs of N_D readings before plotting. In this way the positions of dark and light maxima are plotted instead of the positions of fringe edges. In interferograms measured to date the dark portion of a fringe is approximately flat-topped and narrower than the white.

The graphical process may be replaced entirely by numerical methods e.g., for $j=1, 2 \cdots m$ corresponding to $z_1, z_2 \cdots z_j \cdots z_{m'}$

$$N_u = a_j y + N_{0j} \tag{7}$$

may be evaluated from the undisturbed interferogram. The a_j's may then be averaged or treated in a more sophisticated manner if circumstances require. The result is a linear expression for N_u appropriate to each trace. $\delta=N_D-N_u$ may then be calculated for each N_D value measured in a trace. Interpolation by the Lagrange method will then yield a table of δ *vs* y for each trace. While this method has not been tested as extensively as the graphical one, experience shows it to be not appreciably shorter and somewhat more subject to errors in the interpolation process. If and when the interpolation process can be worked into the ENIAC program and the calculated δ *vs* y values stored suitably for the reduction to follow, then this numerical process will probably be preferable.

Some useful and interesting deductions about the shape of the N_D curve near the shock wave may be

made. From the definition of fringe shift it is clear that,

$$\frac{dN_D}{dy}=\frac{dN_u}{dy}+\frac{d\delta}{dy}=a+\frac{d\delta}{dy}, \tag{8}$$

as N_u is linear with y. Now if we assume that density $(\rho-\rho_0)_u$ is approximately constant within a small distance of the shock wave and equal to the jump ρ_2 at the shock, then for $0.95r_N\leq y\leq r_N$

$$\delta=\frac{2K}{\lambda^*}\int_y^{r_N}\rho_2 d(r_N^2-y^2)^{\frac{1}{2}}=\frac{2K\rho_2}{\lambda^*}(r_N^2-y^2)^{\frac{1}{2}}. \tag{9}$$

Differentiating (9) and substituting into (8) gives,

$$\frac{dN_D}{dy}=a-\frac{2K\rho_2}{\lambda^*}\frac{y}{(r_N^2-y^2)^{\frac{1}{2}}}, \tag{10}$$

from which we see that with fringe numbers chosen so that $\delta>0$ just inside the shock and $a>0$, the N_D curve lies above N_u, is nearly parallel to it for small y and has a continuously turning tangent which passes from $+a$ to $-\infty$ as $y\to r_N$. For negative values of y i.e., the corresponding N_D curve below the axis, the slope at r_N is $+\infty$ and decreases to $+a$ as $y\to 0$. Since our analysis may be expected to hold only in a very narrow region near r_N we inquire how much of the turning of the tangent may occur in this region. It then appears that for a projectile like the 30° half-angle cone at $M=1.77$ (see Fig. 6 where $a\cong 30$ fringes/cm, $\rho/\rho_1=1.8$, $2K\rho_2/\lambda^*=7.89$, 1/cm) the N_D curve has zero slope at $0.97r_N$ and a slope equal to $-a$ at $0.992r_N$. Thus the N_D curve above the axis has a rather sharp hook near the shock wave. If $2K\rho_2/\lambda^*$ remains constant, an increase in a will compress the region from slope 0 to $-\infty$ into even narrower compass. This indicates that an increase in number of fringes per centimeter will not make possible a comparable increase in the amount of data obtainable in the region of sharp turning. More precisely, if we set (10) equal to zero and regard r_N-y_0 as small, there results

$$(r_N-y_0)/r_N\doteq\text{const}/a^2. \tag{11}$$

Doubling the N_u slope decreases the width of the hooked region by 4; so that half as much data can be obtained as before.

Returning to Eq. (9) and setting $y/r_N=f$ we find

$$\delta=\frac{2K\rho_2 r_N}{\lambda^*}(1-f^2)^{\frac{1}{2}}. \tag{12}$$

A plot of δ *vs* $(1-f^2)^{\frac{1}{2}}$ should yield a straight line with slope $2K\rho_2 r_N/\lambda^*$ from which ρ_2 or ρ_2' can be deduced. If a reasonably good straight line can be obtained, then all δ-values near the shock may be assumed to lie on the extrapolated part of the line between the point (0, 0) and the first experimental point obtained. Thus a method is available for completing the δ-curve and the N_D curve near the shock wave in a manner consistent with the assumption of a step discontinuity at the shock. (N.B. This method has not been employed in the reductions reported below. Preliminary results indicate enough improvement to justify including the derivation.)

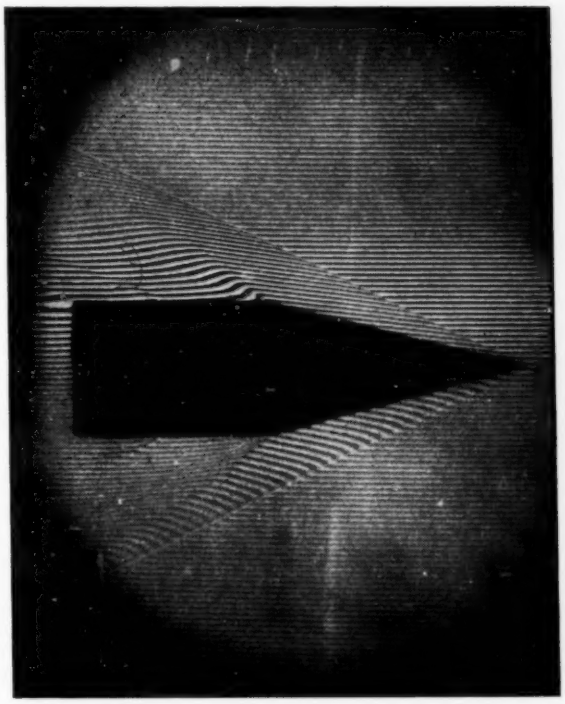


FIG. 2. Interferogram of 15° half-angle, cal 30 cone cylinder at $M=2.7$, $p=1$ atmos.

As in none of the cases treated so far has there been enough data near the shock wave to establish clearly the shape of the N_D curve by experimental means alone, some considerations such as given by this approximate treatment are helpful. In the absence of such information considerable errors may be made in sketching in the shock wave end of the curve.

V. INTERFEROGRAM REDUCTIONS

In Fig. 1 we have a "single fringe" interferogram of a cal 30, 15° half-angle cone-cylinder at $M=2.7$ and $p=1$ atmos. It was taken with the interferometer adjusted for a "single fringe" i.e., with the two beams practically parallel. Under these circumstances the fringes formed in the disturbance are loci of constant fringe shift. Yaw of the projectile is revealed by examination of the shoulder region on each side of the projectile.

Figure 2 shows one of the first interferograms considered suitable for reduction.[12] This record of a cal 30, 15° cone cylinder at $M=2.7$, $p=$atmos indicates relatively poor "stoppage" of the projectile and contrast of the fringes. With undisturbed fringes horizontal, the

[12] F. D. Bennett and W. C. Carter, Phys. Rev. **76**, 880 (A)1949.

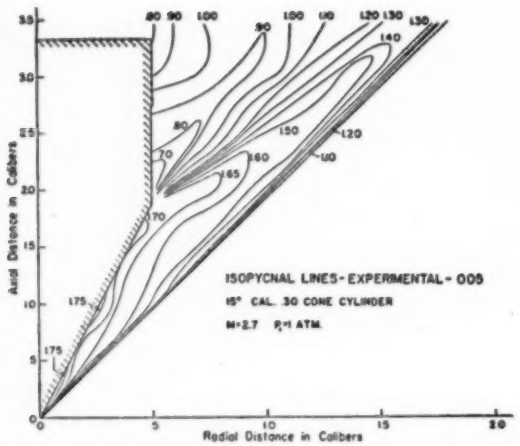

ISOPYCNAL LINES - EXPERIMENTAL - 005
15° CAL. 30 CONE CYLINDER
M=2.7 P₀=1 ATM.
Axial Distance in Calibers
Radial Distance in Calibers

FIG. 3. Isopycnal chart (ρ/ρ_0) for lower half of Fig. 2. $\rho_0 = 1.22$ g/liter.

symmetry of the disturbance is not apparent because all fringes are initially shifted in the same direction. This results in crowding of the fringes below and a separation of the fringes above the centerline. The data for the lower half, because it contains more measurable points, is therefore somewhat better than that of the upper.

Figures 3 and 4 show isopycnal charts (ρ/ρ_0) for the lower and upper halves respectively of Fig. 2. Slight yaw made it impossible to regard the two halves as exactly similar, average the data and thus present one isopycnal chart; yet the agreement is close enough so that either chart may be taken as representative of the general features of the density field. Figure 5 is a theoretical isopycnal chart for the same flow.[13]

In general, comparison of Figs. 3 and 4 with the theoretical chart of Fig. 5 reveals: (a) at the shock wave

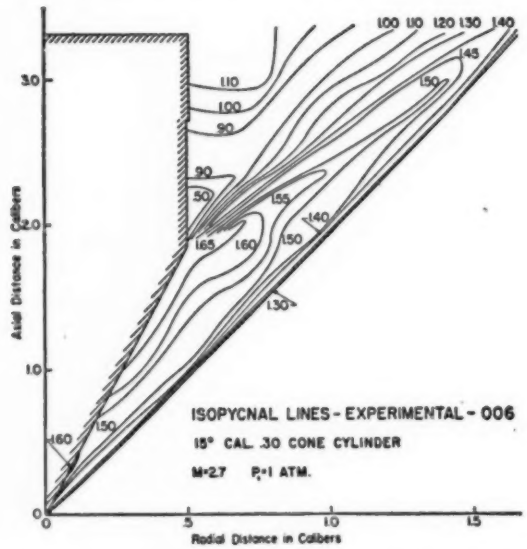

ISOPYCNAL LINES - EXPERIMENTAL - 006
15° CAL. 30 CONE CYLINDER
M=2.7 P₀=1 ATM.
Axial Distance in Calibers
Radial Distance in Calibers

FIG. 4. Isopycnal chart (ρ/ρ_0) for upper half of Fig. 2. $\rho_0 = 1.22$ g/liter.

[13] Clippinger, Giese, and Carter, BRL Report No. 729 and No. 730. The theoretical isopycnal charts were computed from the flow calculations given in these reports.

the jump of 1.30–1.40 compares with 1.39 theoretical, (b) the rise to 1.65–1.70 at the cone compares with 1.60 theoretical, (c) in the expansion at the shoulder 0.80–1.00 compares with 0.86–89 about 2.5 calibers from the nose. In the experimental (ρ/ρ_0) charts lack of sharpness of the shock wave is noticeable. (ρ/ρ_0) values from 1.10–1.40 occur close together instead of an unambiguous 1.39 discontinuity as in the theoretical chart. This smearing out of the shock wave is believed to be caused mainly by the difficulty discussed in connection with Eq. (11) *viz.* the trouble encountered in obtaining a correct fringe shift curve within the last 3–5 percent of the maximum radius r_N. In obtaining the fringe-shift data for these isopycnals and for most of those to follow, a graphical sketching method was used to complete the N_D curve near the shock wave. The results, as will be seen, are better for strong shocks than for

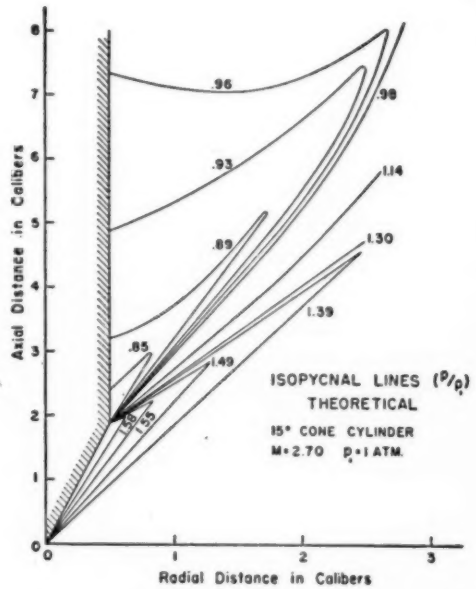


FIG. 5. Calculated isopycnals for projectile of Fig. 2.

weak ones. Where sufficient data can be obtained the method of Sec. IV is believed to be preferable.

In Fig. 6 we see the interferogram of a 30° half-angle, cal 30 projectile at $M = 1.77$ and $p = 1$ atmos. Figures 7 and 8 show the experimental and theoretical isopycnals for this projectile, respectively. In this firing the apparent yaw is practically undetectable, and the base of the projectile is sharp and square except for a slight bulge due to engraving of the driving band by the rifling. Thus the actual yaw may be assumed to be negligible. This assumption is validated by the agreement between fringe-shift measurements for the two halves. The data was averaged and reduced to the single isopycnal chart of Fig. 7.

As before, the "conical" region deviates from conicity to the considerable extent of showing islands near the cone instead of straight lines. The theoretical chart was calculated for $M = 1.70$ while the actual isopycnals

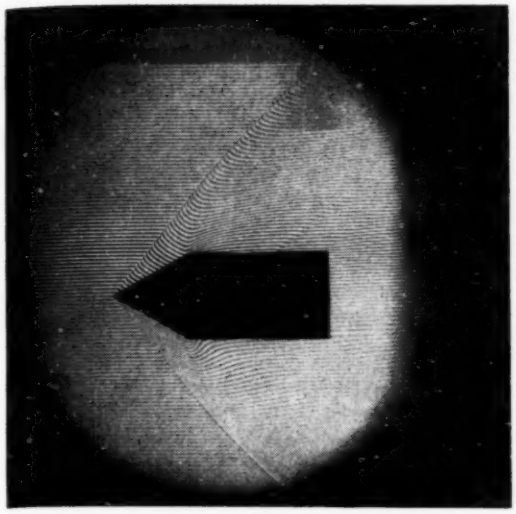


FIG. 6. Interferogram of 30° half-angle, cal 30 cone cylinder at $M=1.77$, $p=1$ atmos $\rho_0=1.19$.

represent $M=1.77$. For this reason the experimental results are all somewhat higher than the theoretical. The jump at the shock appears to be about 1.70 as compared with 1.66 theoretical although the experimental values from 1.70 to 1.80 lie abnormally close to the shock wave. Near the cone values of 1.90–1.92 are reached as compared with 1.86–1.89 of the theoretical plot. The expansion regions are closely comparable; while the 0.65–0.70 "ears" just behind the expansion compare reasonably well except for a reduction in size in the experimental plot.

Figure 6 is directly comparable to the $M=1.7$ 30° half-angle cone-cylinder of Ladenburg, *et al.* (see their Figs. 14 and 15, page 1370) and indeed was produced for purpose of comparison with their results. The two interferograms are remarkably similar especially in the region of the cone. For traces back of the shoulder at points near the shock wave the effects of nonuniformity of the jet are evident in Ladenburg's picture. Here the fringes are much less regular than in our Fig. 6.

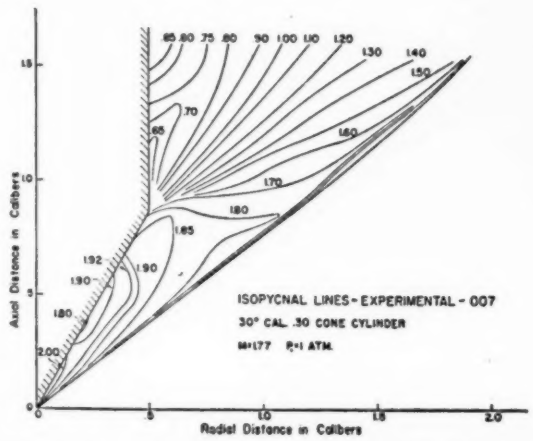


FIG. 7. Isopycnal chart (ρ/ρ_0) for projectile of Fig. 6.

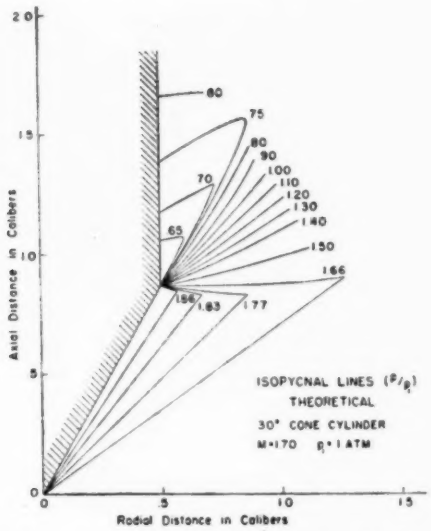


FIG. 8. Theoretical isopycnals for projectile of Fig. 6.

In order to compare the isopycnal charts it is necessary to divide Ladenburg's results in the disturbed region by ρ_0, in this case about 2.15. If this is done, it will be seen that our results are a little higher throughout—as is to be expected from our slightly higher Mach number. Near the cone tip Ladenburg shows a deviation from conical flow similar to ours.[14] Throughout the chart

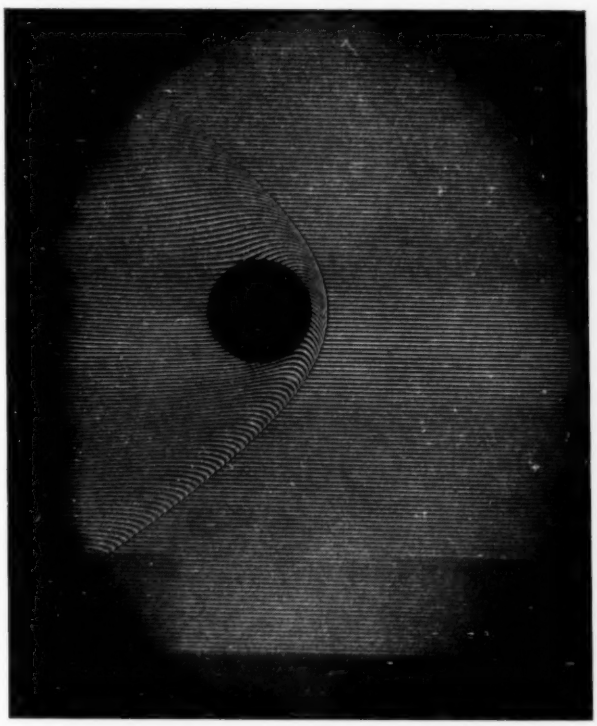


FIG. 9. Interferogram of cal 0.25 sphere at $M=2.0$, $p=1$ atmos.

[14] It may be conjectured that the overshoot caused by nonzero derivative of $(\rho-\rho_0)_{rad}$ at the shock may be partially responsible for the deviation near the tip. This is a region where accumulation of random errors is important as shown in Sec. VI. As the two effects are not separated here no definite conclusion may be drawn.

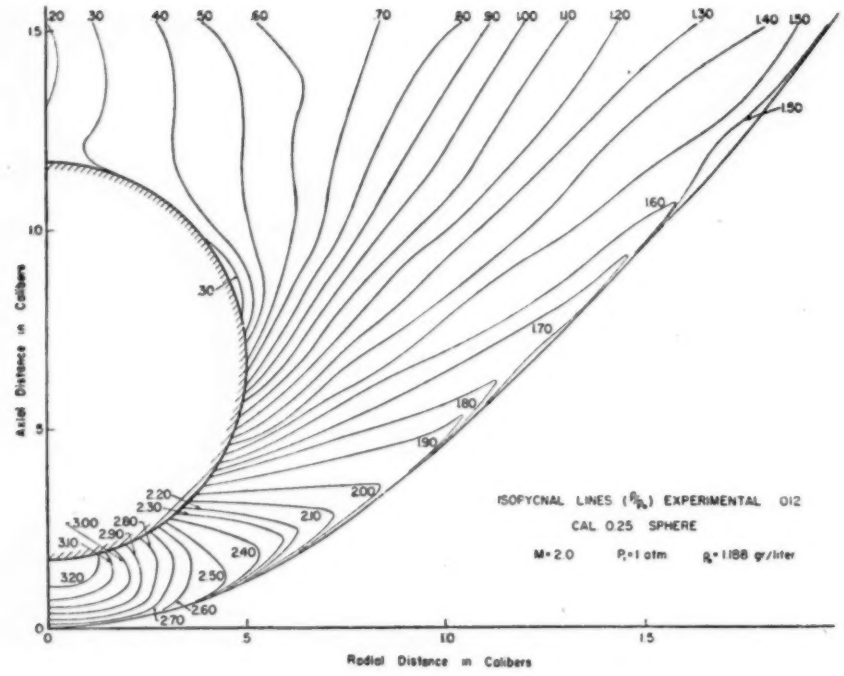

ISOPYCNAL LINES (ρ/ρ_0) EXPERIMENTAL 012
CAL 0.25 SPHERE
M=2.0 P_0=1 atm ρ_0=1188 gr/liter
Axial Distance in Calibers
Radial Distance in Calibers

FIG. 10. Isopycnal chart for cal 0.25 sphere of Fig. 8.

there is agreement of (ρ/ρ_0) values to within about 6 percent. There is a tendency for rise in density from the shock wave conical region to be more abrupt than theory predicts in both Ladenburg's and our results. For this reason the agreement of the experimental plots is somewhat better (about 4 percent less deviation) than agreement between our experimental chart and the theoretical.

Figure 9 shows an interferogram of a cal 0.25 sphere at $M=2.0$ and $p=1$ atmos. This picture may be compared with the sphere of Ladenburg, Winckler, and Van Voorhis,[1] see p. 1376, and the sphere interferograms recently reported by Gooderum and Wood.[15] Although these authors report data for Mach numbers 1.70 and 1.30, 1.62, respectively, the interferograms are all very similar and comparable in so far as main features are concerned. Here again one sees that the free flight interferogram is free of the irregularities caused by lack of homogeneity near the edge of the jet.

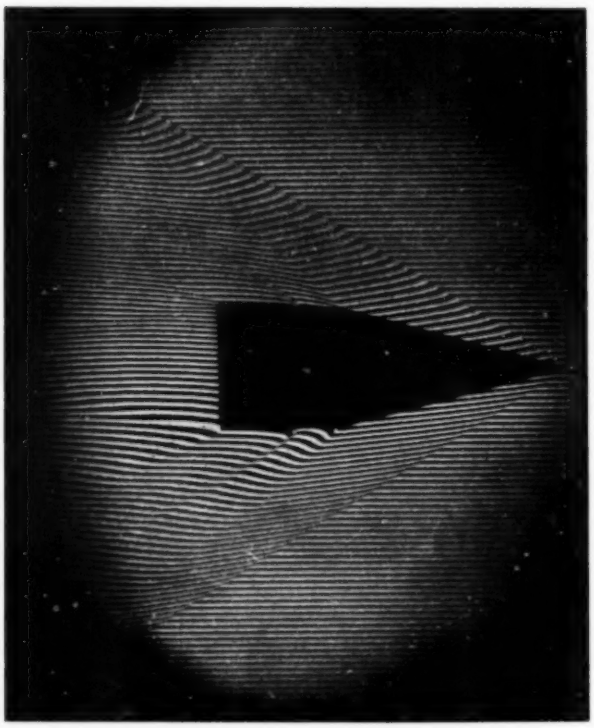


FIG. 11. Interferogram of 15° half-angle, cal 30 cone cylinder at $M=2.16$, $p=1$ atmos.

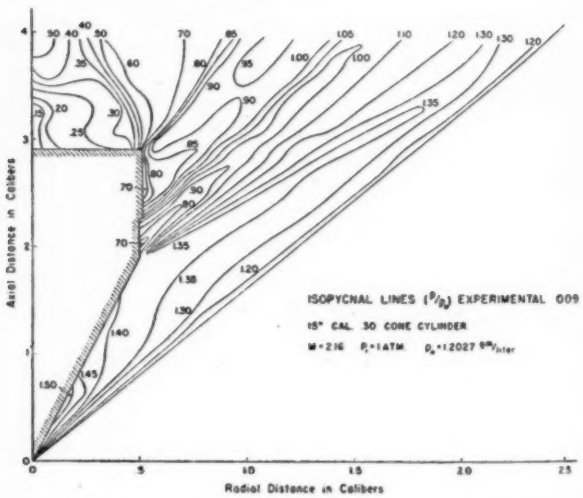

ISOPYCNAL LINES (ρ/ρ_0) EXPERIMENTAL 009
15° CAL. 30 CONE CYLINDER
M=2.16 P_0=1 ATM. ρ_0=.2027 gm/liter
Axial Distance in Calibers
Radial Distance in Calibers

FIG. 12. Isopycnal chart (ρ/ρ_0) for upper half of Fig. 11.

[15] P. B. Gooderum and G. P. Wood, NACA TN 2173 (August, 1950).

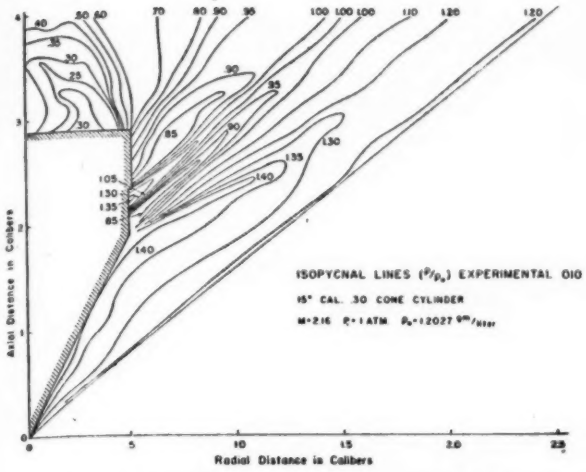

ISOPYCNAL LINES (ρ/ρ_0) EXPERIMENTAL 010
15° CAL. 30 CONE CYLINDER
Axial Distance in Calibers
Radial Distance in Calibers

FIG. 13. Recalculated isopycnals for upper half Fig. 11. Smoothing of data improved.

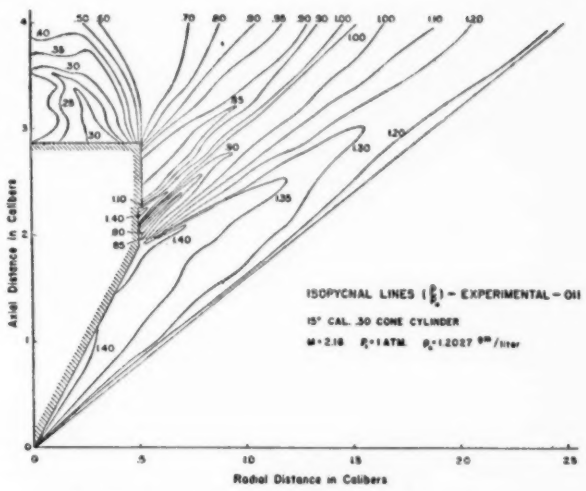

ISOPYCNAL LINES ($\frac{\rho}{\rho_0}$) - EXPERIMENTAL - 011
15° CAL. 30 CONE CYLINDER
Axial Distance in Calibers
Radial Distance in Calibers

FIG. 14. Isopycnal chart for lower half of Fig. 11.

Figure 10 is the isopycnal chart for the cal 0.25 sphere. Comparison with the reductions of the papers cited above reveals some pinching of the isopycnals found by Ladenburg near the shock wave—probably because of the inhomogeneity of the flow in front of the shock. In the reductions of Gooderum and Wood this effect is not noticeable because these authors did not extend their curves to the shock wave. They did not carry out a "deshocking" process at the shock wave to account for the density jump there, and consequently, obtained an overshoot just inside the shock wave; thus were unable to rely on their values near the shock wave. Otherwise the isopycnals are very similar in the three cases. From the experimental (ρ/ρ_0) values in Fig. 10 the stagnation pressure on the front surface of the sphere has been calculated. Deviation from the theoretical value was about 4 percent. Similarly, calculation of the density jump at various points along the shock wave shows good agreement with the experimental values.

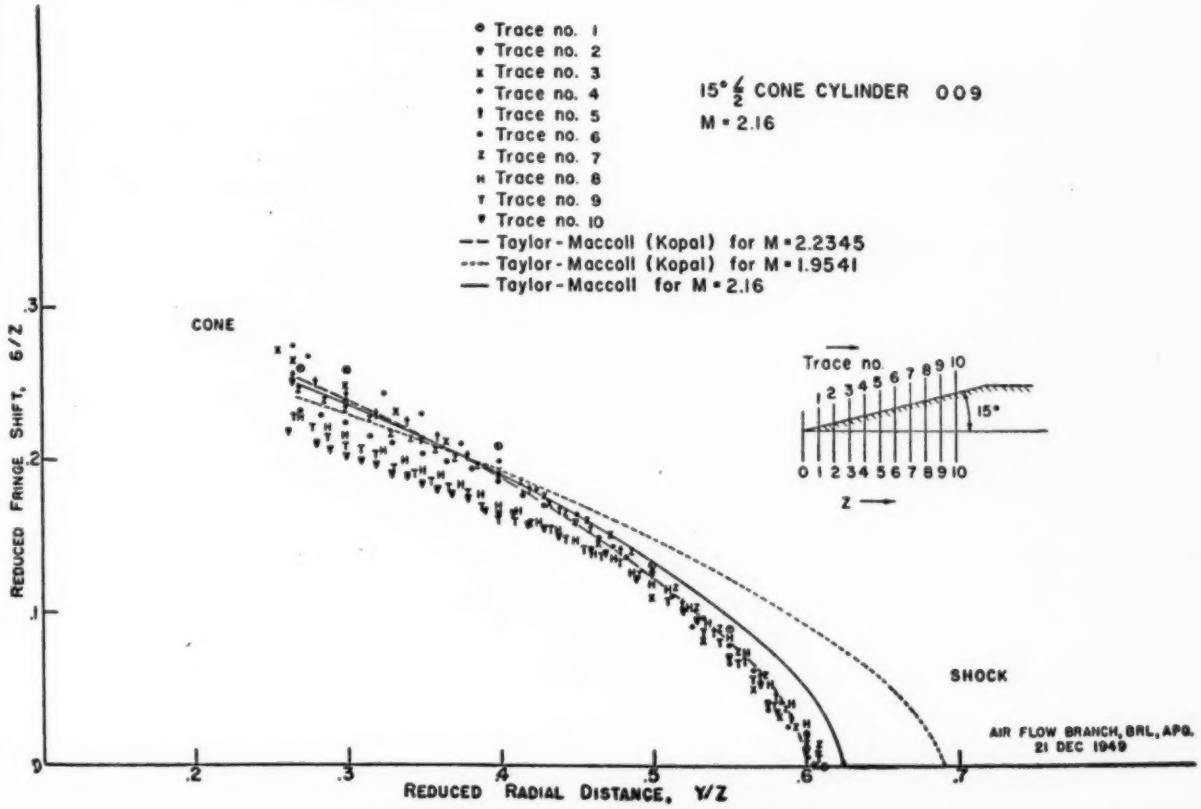


FIG. 15. Conical flow test for fringe shift data of Fig. 14.

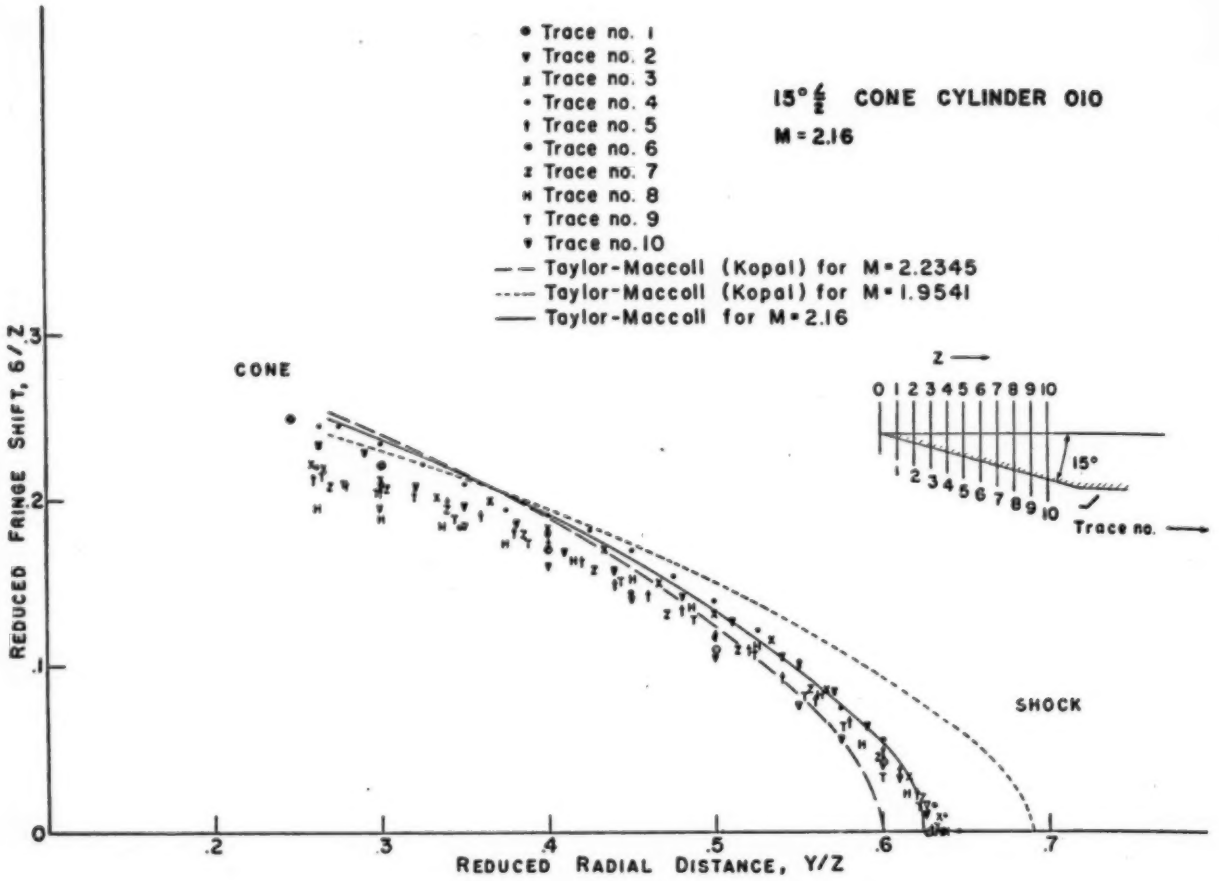


FIG. 16. Conical flow test for fringe shift data of Fig. 12.

In Fig. 11 we have an interferogram of a cal 30 15° half-angle cone cylinder at $M=2.16$ and $p=1$ atmos. Figures 12 and 13 give two successive reductions of the data for the upper half, i.e., half with the most fringes, of Fig. 11. To obtain Fig. 13 the fringe shift data for Fig. 12 was re-examined, the smoothing was checked against the interferogram to see that no actual irregularities were arbitrarily smoothed out. Except for local changes of (ρ/ρ_0), about 5 percent or less in some regions, the two charts are nearly identical.

Figure 14 shows the isopycnal chart for the lower half of the Fig. 11 interferogram. Here, the influence of slight yaw is seen in the (ρ/ρ_0) values near the cone. These isopycnals are more curved and as much as 10 percent higher than those of the upper half. The comparison is otherwise very close except in the base region near the axis. Lack of symmetry with the base region of Fig. 13 is noticeable. This discrepancy is believed caused primarily by the accumulation of random errors in the calculation as is shown in Sec. VI.

In order to test the consistency of the reduction process, i.e., to investigate whether large errors not present in the fringe shift data may be introduced by the process into the $(\rho-\rho_0)$ values, conical flow plots[8] have been made for both the fringe shift data and the calculated density ratios for both halves of the previous interferogram. These plots are shown in Figs. 15–18.

The plots of (ρ/ρ_0) in Figs. 17–18 are not strictly comparable with those of δ/z given in Figs. 15–16. The quantity $(\rho-\rho_0)$ is more directly related to δ/z than (ρ/ρ_0). One is able to show easily that percent error in $(\rho-\rho_0)=(\rho/\rho_0)(\rho/\rho_0-1)^{-1}\times$percent error in (ρ/ρ_0). For the (ρ/ρ_0) values with which we are concerned this means that percent error in $(\rho-\rho_0)$ is much larger than percent error in (ρ/ρ_0).

To facilitate comparison of the two types of conicity test Table I has been prepared from the data represented in Figs. 15 and 17.

At several y/z values of Fig. 17, percent $\Delta(\rho/\rho_0)$ has been calculated from the maximum spread of the (ρ/ρ_0) values divided by a mean value (ρ/ρ_0). This has been converted to percent $\Delta(\rho-\rho_0)$ as indicated above; finally percent $\Delta(\delta/z)$ has been found in a similar manner from the graphical values of Fig. 15. The 4th and 5th rows of the table are to be compared directly.

At the shock wave $(y/z=0.6)$ the spread is largest percentage wise but the agreement is fairly good. At the next point the calculated values still show a slight advantage, but thereafter have larger percentage spread than the primordial fringe shift values. The maximum

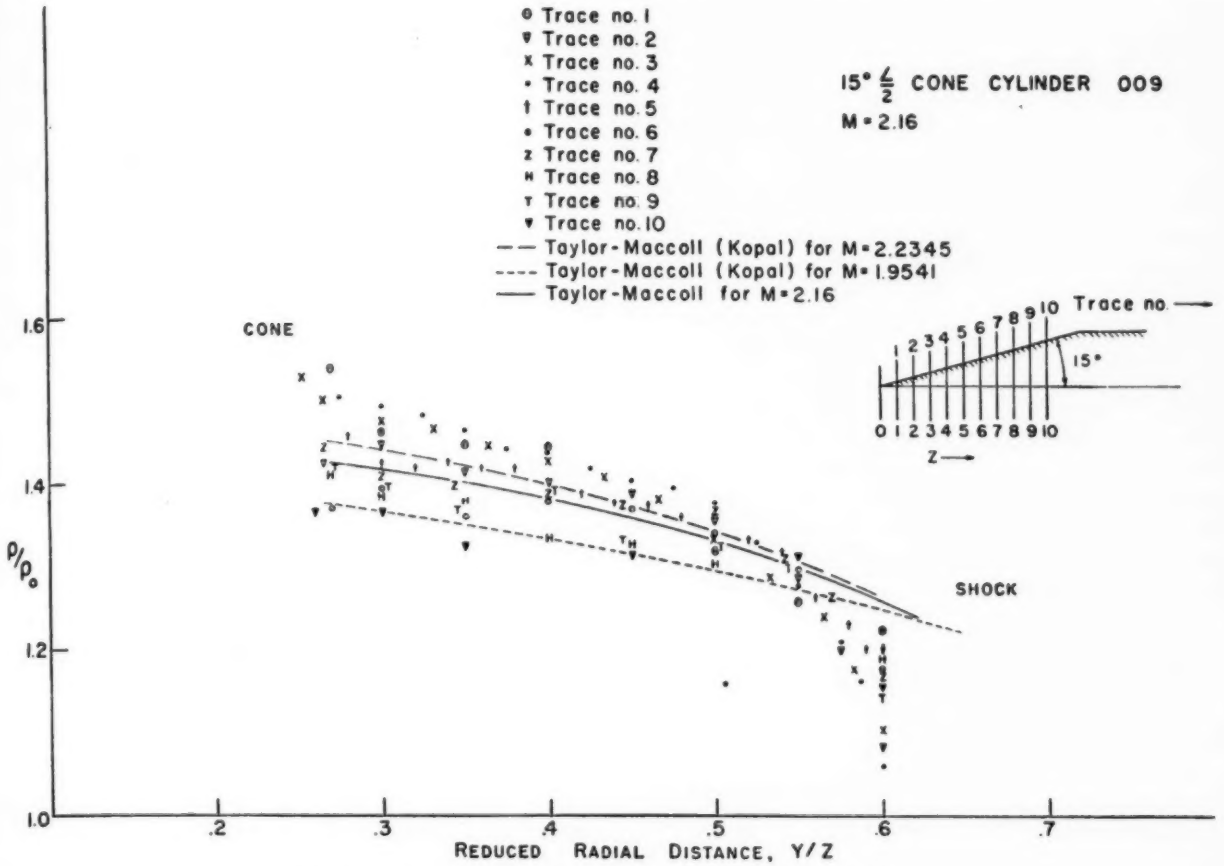


FIG. 17. Conical flow test of (ρ/ρ_0) values plotted in Fig. 14.

accumulated error, i.e., difference in percentage spread, is about 13 percent.

It is clear that in the conical region at least, and probably elsewhere, the calculated (ρ/ρ_0) values have deteriorated considerably error wise, from the data from which they arise. The cause of this phenomenon is not completely understood at present. The spread at the shock wave may be considerably reduced by the procedure of Sec IV, which will improve both the (δ/z) and (ρ/ρ_0) values. The accumulated error is not throught to be significant here in view of the large percentage error of both quantities. The increase of accumulated error toward the body may be in large part because of random error which, as will be seen in Sec VI, increases toward the axis of flow.

VI. SOURCES OF ERROR IN THE REDUCTION PROCESS

(a) To investigate the propagation of random error in δ into the reduction process we refer to Eq. (5). If we divide the disturbance into N equal zones of width w beginning at the axis, we may set

$$r_j = jw \tag{13a}$$

$$\Gamma = 2\lambda^*/\pi K\rho_0 \tag{13b}$$

$$A_{ji} = \frac{1}{((j+1)^2-i^2)^{\frac{1}{2}}+(j^2-i^2)^{\frac{1}{2}}}. \tag{14}$$

Then if e_j^k represent the random errors in fringe shift at the jth zone for the kth member of a sequence of m similar measurements, the error in $(\rho/\rho_0)_i$ is

$$\Delta_i^k = -\frac{\Gamma}{w}\sum_{j=i}^{N-1}(e_{j+1}^k - e_j^k)A_{ji}. \tag{15}$$

If we now assume a normal distribution law for the e_j^k's it may be shown by straightforward calculation[16]

TABLE I.

(1) y/z	0.6	0.55	0.50	0.45	0.40	0.35	0.30	0.26
(2) $\%\Delta(\rho/\rho_0)$	14.4	4.2	5.7	6.6	7.8	7.5 10.0	9.0	11.0
(3) $\dfrac{\rho/\rho_0}{\rho/\rho_0-1}$	8.15	4.46	3.94	3.78	3.56	3.44 3.56	3.32	3.17
(4) $\%\Delta(\rho-\rho_0)$	117	19	22	25	28	26[a] 36	30	35
(5) $\%\Delta(\delta/z)$	130	26	17	14	26	22	25	23
(6) Accumulated % error	−13	−7	5	11	3	4 14	5	12

[a] The smaller value omits the low point of Fig. 17.

[16] A. G. Worthing and J. Geffner, *Treatment of Experimental Data* (John Wiley and Sons, Inc., New York, 1943), pp. 208–209.

that Σ_i, the rms error in $(\rho/\rho_0)_i$, is related to σ_j, the rms error in fringe shift by,

$$\Sigma_i=\frac{\Gamma}{w}\sum_{j=i}^{N-1}\sigma_j^2(A_{j-1i}-A_{ji})^2.^{17} \qquad (16)$$

This is a not unexpected result, but does not seem to have been obtained before for this reduction problem. The coefficients A_{ji} are readily available since the publication of Gooderum and Wood[15] who give a table for 50 zones. A 60-zone table has been in use at this laboratory since 1948 but has not been published as interest turned to machine methods of computation. With A_{ji} values and σ_j's given (16) gives a readily computable expression for the rms error in $(\rho/\rho_0)_i$; however, certain simplifications occur when the accuracy of measurement is assumed constant across the plate i.e., $\sigma_j=\sigma=$const. With this condition a very good bound on Σ_i may be obtained sufficient for most purposes of calculation.

Thus (16) becomes

$$\Sigma_i=\frac{\Gamma}{w}\sigma\Big(\sum_{j=i}^{N-1}(A_{j-1i}-A_{ji})^2\Big)^{\frac{1}{2}}. \qquad (17)$$

For $j\neq i$, since $A_{j-1i}>A_{ji}$

$$\Sigma_i\leq\frac{\Gamma}{w}\sigma\Big(A_{ii}^2+\sum_{i+1}^{N-1}(A_{j-1i}^2-A_{ji}^2)\Big)^{\frac{1}{2}},$$

in which the summation involves a telescopic series whence,

$$\Sigma_i\leq\frac{\Gamma}{w}\sigma(2A_{ii}^2-A_{N-1i}^2)^{\frac{1}{2}}\leq\frac{\sqrt{2}\Gamma\sigma A_{ii}}{w}. \qquad (18)$$

For any reasonable number of zones the A_{N-1i}^2 term is very small; so its absence little affects the bound.

From the definition, $A_{ii}=1/(2i+1)^{\frac{1}{2}}$ which combined with $r_i=iw$ (18) yields

$$\Sigma_i\leq\Gamma\sigma\left[\frac{2}{w(2r_i+w)}\right]^{\frac{1}{2}}\leq\frac{\Gamma\sigma}{wr_i}. \qquad (19)$$

For a fixed value of r_i, decrease in zone width by the factor K increases the number of zones by the same factor, but increases the bound (19) on rms error by only $K^{\frac{1}{2}}$.

From (19) it is clear that points near the axis will have the larger random errors; and, in fact, the axial

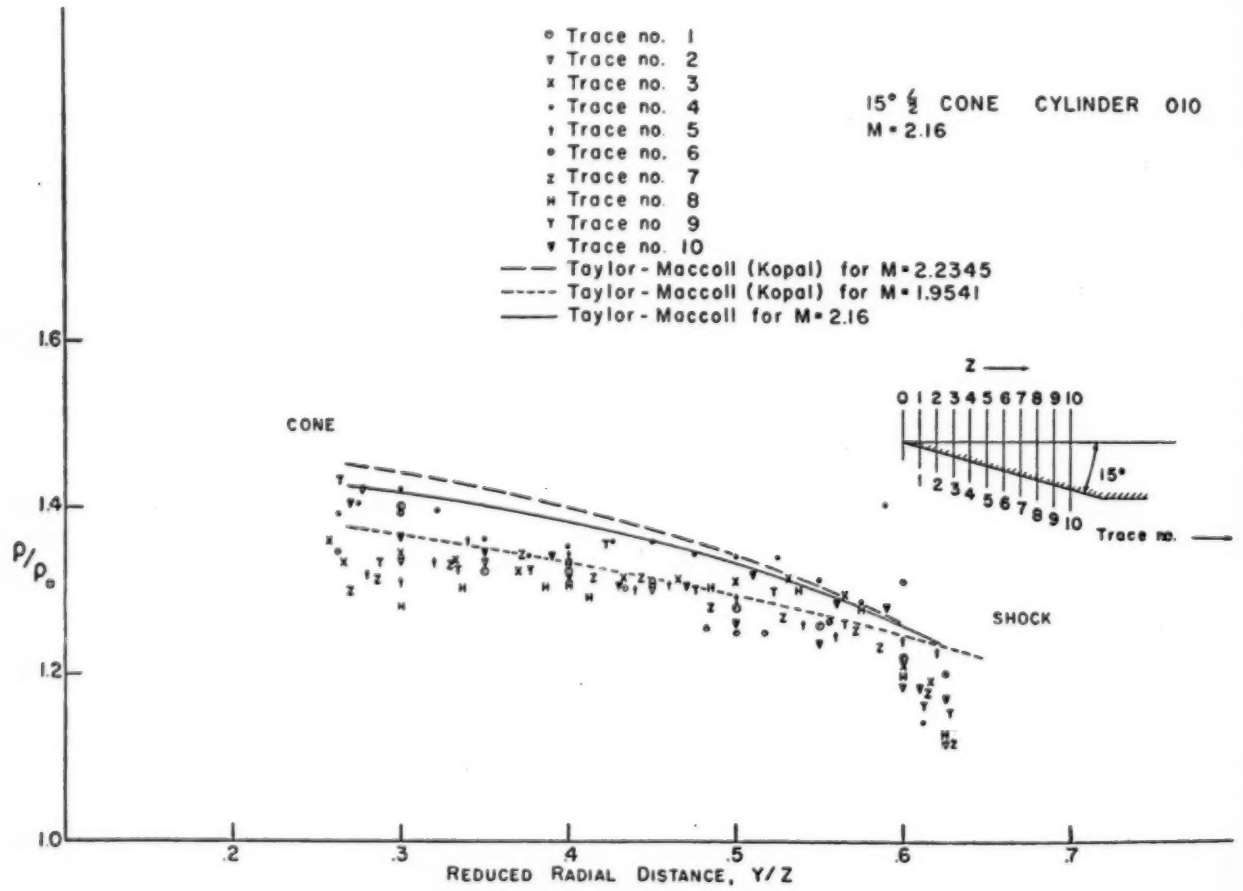


FIG. 18. Conical flow test of (ρ/ρ_0) values plotted in Fig. 12.

[17] Note the convention that $A_{i-1i}=0$.

point of each trace $r_i=0$ will be the worst. As zone size decreases the minimum radius for a given bound on the error moves away from the axis.

Thus if minimum error m be assigned and $r_i \gg w$,

$$m=\frac{\sqrt{2}\Gamma\sigma}{(2r_iw+w^2)^{\frac{1}{2}}}$$

or

$$r_i=\frac{\Gamma^2\sigma^2}{wm^2}-\frac{w}{2}\doteq\frac{\Gamma^2\sigma^2}{wm^2}.$$

We have here an indication why the density values in the base regions of Figs. 13 and 14 fail to agree. The region near the axis where the isopycnals should join is just that region where random errors are most severe; furthermore, ρ/ρ_0 takes on its smallest values here with the consequence that large percentage errors may occur.

To form an idea of the magnitude of the random errors in $(\rho/\rho_0)_i$ let us substitute some representative values into (19) and compare the results with the corresponding isopycnal chart. The comparison will not be strict because two zones sizes have been used in actual reductions ($w=0.01$ cm over most of the trace and $w=0.005$ cm near the shock) while (19) is based on constant w; nevertheless, the magnitudes of errors are very closely representative. For $\lambda=5461$A and $\rho_0=1.20$ g/liter, $\Gamma=0.128$. We use $w=0.01$ cm and $\sigma=0.01$. This estimate of the rms random error in fringe width of 0.01 fringe is the lowest possible with the present method of determining δ. It is felt that this accuracy has been obtained in some of the better measurements but is by no means general for all of the reductions reported here. We find

$$\Sigma_i\leq 0.013/(r_i+w/2)^{\frac{1}{2}}.$$

Table II shows the resulting values. Since the Σ_i values represent numerical errors it is necessary to know the $(\rho/\rho_0)_i$ value at a particular radius in order to estimate percent error. Referring now to Fig. 13 as an example for which $\sigma=0.01$ should hold, and remembering that $\frac{1}{2}$ cal$=0.381$ cm we see that in the tip region at about 0.1 cm $(\rho/\rho_0)\geq 1.20$ and the random error will be less than about 3.5 percent (percent $\Delta(\rho-\rho_0)\leq 20$ percent). The isopycnals are abnormally straight in this region because (1) all are arbitrarily faired through the apex of the cone, and (2) the first trace occurs about 1 mm ($\sim\frac{1}{3}$ body radius) back of the tip.

As the shoulder $r_i\geq 0.381$ and the percentage error on the 1.40 isopycnal is 1.5 percent (percent $\Delta(\rho-\rho_0)\leq 5.2$). At twice the body radius the error on the ~ 1.33 isopycnal is ~ 1.2 percent (percent $\Delta(\rho/\rho_0)\leq 4.5$). Deviations from a straight isopycnal in this region seem to be of the order of 2–3 percent. This fact suggest either that the σ value is low by a factor of about 2, or that systematic errors in the measured δ-values caused adjacent $(\rho/\rho_0)_i$ traces to be "radially displaced" relative to each other in such a way as to simulate a random variation of the isopycnal from a smooth average curve. Inspection of the trace plots of $(\rho/\rho_0)_i$ favors the latter conclusion as scatter of the points from an average curve appears to be less than 1 percent almost everywhere on the trace. We then conclude that for most of the area outside the cylinder in the region bounded by the $(\rho/\rho)_i=1.00$ isopycnal, the random error should be ≤ 2.1 percent, usually considerably less. For lower values of $(\rho/\rho_0)_i$ the percentage error is correspondingly higher.

Table II.

r_i (cm)	0	0.1	0.2	0.381	0.5	0.8	1	4
Σ_i	0.19	0.04	0.029	0.021	0.018	0.015	0.013	0.006

In the base region at $r_i=0.1$ cm the numerical error is 0.04 but because of the low $(\rho/\rho_0)_i$ values, may be more than 15 percent of the total. At the axis itself the value is probably 50–100 percent in error and completely untrustworthy. Inspection of the $(\rho/\rho_0)_i$ traces at the base bears out this conclusion.

From the percent $\Delta(\rho/\rho_0)$ values given in the immediately preceeding discussion the accumulated error in reduction is of the order 5–8 percent except for traces at small radii near the tip where a 20 percent bound appears. Inspection of the conicity plots shows that at some of the inner y/z values e.g., 0.6 in Fig. 17 the scatter of points on the first 3–4 traces adds materially to the spread between maximum and minimum values represented in Table I. We may then tentatively conclude that accumulated random error is mainly responsible for the deterioration in accuracy because of the reduction; and that, in the example studied here, the tip traces at small radius values contribute most of the error.

(b) To investigate error resulting from approximation to the integral we quote the bound obtained in (I-15), *viz.*

$$E\leq\frac{\pi\Gamma\left|\delta_{uu}(u_N-u)^{\frac{1}{2}}\right|_{\max}r_N^2}{N}, \qquad (20)$$

which bounds the approximation error for every zone in a trace divided into N equal intervals in r space ($r^2=u$). Evaluation of the max quantity may be done approximately if tables of reduced fringe shift against zone number j are available.

Calculations show that the bound (20) is too crude to be of much value in estimating actual N values. Perhaps the most that can be said is that to keep (20) constant from trace to trace requires $N\propto r_N^2$, other parameters remaining the same

(c) The question of refraction near the shock wave cannot be satisfactorily treated at present. If we idealize the disturbance as a cone of index $n_2>n_1$ with n_1 the undisturbed index of refraction, it is not hard to show that the tangent ray proceeding normal to the cone

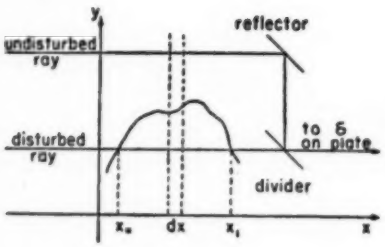


FIG. 19. Diagram of light ray passing through the disturbance.

axis will suffer a total internal reflection. Indeed for cone angle $\pi/6$, $n_2=1.000544$ and $n_1=1.000277$ the first ray to emerge corresponds to a ray making 0.01 radian less than $\pi/2$ with the radius from the axis to the point of incidence.[18]

The conical disturbance may be treated approximately by assuming that in the vicinity of the tangential ray, refraction occurs in the plane defined by this ray and the cone-normal at the point of tangency. With the additional assumption that the intersection of cone and plane may be replaced by the osculating circle, R. F. Clippinger[19] has obtained an approximation for the fringe shift with refraction. At the shock where unrefracted fringe shift is zero, a negative value is obtained. Thus fringe shift with refraction changes from minus to plus values in a very narrow region near the shock. A similar result is obtained if we replace the cone by a cylinder of radius equal to that of the cone at the point of tangency of the incident ray. Incidentally, the treatment of Gooderum and Wood[15] ignores the effect of camera lens in recombining interfering rays and is thus incapable of showing the negative fringe shift.

Studies of fringe shift near the shock wave using the technique of Sec. IV, show that data near the shock usually lies below the expected line. This is interpreted at present as evidence that the negative shift occurs, although unresolved by present methods, and that on account of refraction the first observable points are slightly low.

(d) Yaw can lead to serious departures from axial symmetry and thus cause errors in the reductions based upon an assumption of strict axisymmetry. In the preceding interferogram reductions, e.g., Figs. 3 and 4, discrepancies as high as 5–10 percent have been noted between the two halves of an interferogram reduced as if each half represented an axisymmetric flow. While some of this discrepancy is doubtless because of accumulated random error of the reduction process, the greater part is believed to represent the effects of yaw. Indeed the presence of higher density ratios on the half with positive angle of attack confirms this view.

The reduction of interferograms of flows about slightly yawed bodies has been discussed by several authors,[20] but has never been seriously attempted to the knowledge of the present writers. Dr. J. H. Giese of this Laboratory has developed the equations for fringe shift about a slightly yawing body of revolution. Assuming that the plane of symmetry can be found and that each cross section of the flow is nearly circular, equations for the zeroth and first-order perturbation of the density function can be found. Both equations may be solved by the Abel method; so that in principle a treatment of the slightly yawing projectile is possible.

(e) Finally, brief mention may be made of the effect of errors in the various constants on the accuracy of the reduction. Since most of these enter through the constant term multiplying the sum in Eq. (5), the effect will be a systematic error in the calculated density values. It is believed that the total effect on the constant can be held to less than 1 percent.

APPENDIX I

Theory of Interferogram Reduction

1. Introduction

We shall recall a few of the basic assumptions of geometrical and physical optics: (1) in a homogeneous, isotropic medium a beam of light may be analyzed into rays which are represented either as vectors or straight lines in three-space, (2) a monochromatic ray of light consists of a sinusoidal disturbance of frequency ν and wavelength λ propagating along the normal to the direction of vibration with velocity v, (3) $\lambda\nu=v$ is satisfied where ν is constant and λ, v are functions of the coordinates, (4) index of refraction n is related to velocity v by $n=c/v$ where c is the velocity of light in vacuum, (5) for gaseous media such as the permanent gases at moderate pressures there is a direct proportion between density increments $\rho-\rho_0$ and index of refraction increments $n-n_0$ thus,

$$n-n_0=K(\rho-\rho_0), \tag{I-1}$$

where K is the Gladstone-Dale constant. Finally, we make a special assumption necessary to simplify the problem of analyzing axisymmetric gas flows *viz.* (6) the refraction of any ray in the disturbed beam is

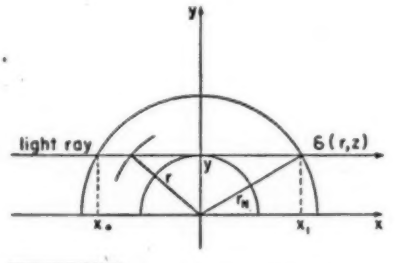


FIG. 20. Axisymmetric disturbance.

[18] We are indebted to Mr. Raymond Sedney for this result. Circumstances may also occur in which the tangent ray will never emerge from the cone.

[19] BRL Report No. 576, Aberdeen Proving Ground, 21 September (1945).

[20] H. Schardin is believed to have discussed this problem at Saint-Louis, France, R. Sauer is reported to have presented equations for such a reduction at a meeting at N.O.L., White Oaks, Maryland, but neither author has published anything on the subject to our knowledge.

negligible.[21,22] With these preliminaries we may now pass to the derivation of the fringe shift equation.

2. Fringe Shift Equation

Consider a general density disturbance in one beam of a Mach-Zehnder interferometer. Coordinate directions are defined by an orthogonal right-hand triad with $\bar{x}$ parallel to the interferometer beam and $\bar{y}$ vertical as in Fig. 19.

Then if the index of refraction function $n(x, y, z)$ has only a finite number of finite discontinuities, the fringe shift $\delta(y, z)$ is given by,

$$\delta(y,z)=\frac{1}{\lambda^*}\int_{x_0}^{x_1}(n-n_0)dx=\frac{K}{\lambda^*}\int_{x_0}^{x_1}(\rho-\rho_0)dx, \quad \text{(I-2)}$$

where ρ_0 and n_0 refer to density and index of refraction, respectively, in the undisturbed beam and λ^* is the vacuum wavelength of the light.

To prove (I-2) we observe that the optical path of the disturbed ray measured in wavelengths is

$$N_d=\int_{x_0}^{x_1}\frac{dx}{\lambda}.$$

The difference N_d-N_0 gives the number of multiples of 2π by which the two rays differ in phase when recombined. This number is just equal to the fringe shift $\delta(y, z)$. (I-2) follows from (I-1) and hypothesis (3).

If $n=n(y, z)$ and (I-1) holds, then for two-dimensional flows,

$$\rho-\rho_0=\frac{\delta\lambda^*}{K(x_1-x_0)}=\frac{\delta\lambda^*}{Kl}. \quad \text{(I-2a)}$$

3. Transformation of the Fringe Shift Equation

The integral Eq. (I-2) may be transformed into a type studied by Abel in case the disturbance is axisymmetric. Assuming that the z axis is the axis of symmetry we see upon reference to Fig. 20 that,

$$\delta(y,z)=\frac{2K}{\lambda^*}\int_0^{x_1}(\rho-\rho_0)dx,$$

and since $r^2=x^2+y^2$ with $y=\text{const}$,

$$\delta(y,z)=\frac{2K}{\lambda^*}\int_y^{r_N}\frac{(\rho-\rho_0)rdr}{(r^2-y^2)^{\frac{1}{2}}}. \quad \text{(I-2b)}$$

If $r^2=v$, $y^2=u$, and $r_N^2=u_N$, we have

$$\delta(u_i,z)=\frac{K}{\lambda^*}\int_{u_i}^{u_N}\frac{(\rho-\rho_0)dv}{(v-u_i)^{\frac{1}{2}}}. \quad \text{(I-2c)}$$

[21] R. F. Clippinger, BRL Report No. 576 (Aberdeen Proving Ground, September 21, 1945).

[22] The problem of correcting for refraction errors in the interferometry of two dimensional flows is studied by R. Ladenburg, Contract No. N7 ONR 399 (December 15, 1949), R. E. Blue, NACA Technical Note 2110 (June 1950), L. De Frate thesis (1950) Massachusetts Institute of Technology, Cambridge, Massachusetts; however since we are principally interested in three-dimensional axisymmetric flows with short path lengths and small density gradients we shall assume (6) to hold.

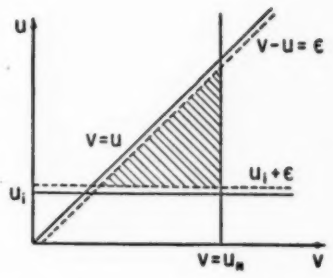


FIG. 21. Region of integration.

4. Solution of the Integral Equation

We employ an adaptation of the treatment given by F. J. Weyl.[4] Assume (1) $\rho(u, z)$ is continuous and has piecewise continuous first and second derivatives in u and z, (2) for $u\geq u_N$, $\rho=\rho_0=\text{const}$, and (3) the Eq. (I-2c). It follows that

$$(\rho-\rho_0)_{u_i}=\frac{-\lambda^*}{\pi K}\int_{u_i}^{u_N}\frac{d\delta/du}{(u-u_i)^{\frac{1}{2}}}du. \quad \text{(I-3)}$$

To prove this, substitute $u_i=u$ in (I-2c), multiply on both sides by $du/(u-u_i)^{\frac{1}{2}}$ and integrate from u_i to u_N; thus

$$\int_{u_i}^{u_N}\frac{\delta(u,z)du}{(u-u_i)^{\frac{1}{2}}}=\frac{K}{\lambda^*}\int_{u_i}^{u_N}\int_{u}^{u_N}\frac{(\rho-\rho_0)dv}{(v-u)^{\frac{1}{2}}}\frac{du}{(u-u_i)^{\frac{1}{2}}}. \quad \text{(I-3a)}$$

The integrand of the double integral has discontinuities at $u=u_i$ and $v=u$. To remove this difficulty consider as in Figure 21 the triangular region R^* in the u, v plane bounded by the lines $v-u=\epsilon$, $u=u_i+\epsilon$, and $v=u_N$, where $0<\epsilon\ll 1$.

If we modify the integral so that the region of integration is R^* rather than the larger triangle R containing R^*, the integrand $(\rho-\rho_0)/[(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}]$ is clearly continuous in and on the boundary of R^*. We may thus apply Dirichlet's formula[23] and interchange the order of integration. The result is

$$\int_{u_i+\epsilon}^{u_N}\int_{u+\epsilon}^{u_N}\frac{(\rho-\rho_0)dvdu}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}=\int_{u_i+2\epsilon}^{u_N}(\rho-\rho_0)dv\int_{u_i}^{v-\epsilon}\frac{du}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}. \quad \text{(I-3b)}$$

Now,

$$\int_{u_i}^{v}\frac{du}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}=\int_{u_i}^{u_i+\epsilon}+\int_{u_i+\epsilon}^{v-\epsilon}+\int_{v-\epsilon}^{v};$$

so

$$\left|\int_{u_i}^{v}\frac{du}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}-\int_{u_i+\epsilon}^{v-\epsilon}\frac{du}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}\right|\leq\left|\int_{u_i}^{u_i+\epsilon}\right|+\left|\int_{v-\epsilon}^{v}\right|.$$

[23] D. V. Widder, *Advanced Calculus* (Prentice-Hall, Inc., New York, 1947), p. 165.

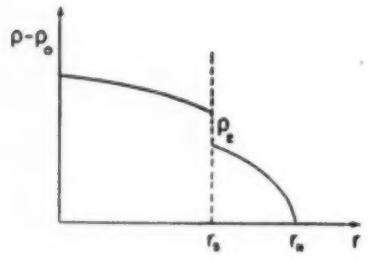


FIG. 22. $\rho-\rho_0$ with step discontinuity.

$$\left|\int_{u_i}^{u_i+\epsilon}\frac{du}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}\right|$$

$$\leq\frac{1}{(v-u_i-\epsilon)^{\frac{1}{2}}}\left|\int_{u_i}^{u_i+\epsilon}\frac{du}{(u-u_i)^{\frac{1}{2}}}\right|=\frac{2\sqrt{\epsilon}}{(v-u_i-\epsilon)^{\frac{1}{2}}}. \quad \text{(I-3c)}$$

Similarly for $\int_{v-\epsilon}^{v}$.

By using these results and No. 113 Peirce, *Table of Integrals*, (I-3c) becomes

$$\left|\pi-\int_{u_i+\epsilon}^{v-\epsilon}\frac{du}{(v-u)^{\frac{1}{2}}(u-u_i)^{\frac{1}{2}}}\right|\leq\frac{4\sqrt{\epsilon}}{(v-u_i-\epsilon)^{\frac{1}{2}}}.$$

Multiply this expression by $K/\lambda^*(\rho-\rho_0)dv$ and integrate within the absolute value sign from $u_i+2\epsilon$ to u_N, and we find

$$\left|\frac{\pi K}{\lambda^*}\int_{u_i+2\epsilon}^{u_N}(\rho-\rho_0)dv-Q(\epsilon)\right|$$

$$\leq\text{const. }(\epsilon)^{\frac{1}{2}}[(u_N-u_i-\epsilon)^{\frac{1}{2}}-(\epsilon)^{\frac{1}{2}}]'$$

where $Q(\epsilon)$ represents the right side of (I-3b). As $\epsilon\to 0$,

$$\frac{\pi K}{\lambda^*}\int_{u_i}^{u_N}(\rho-\rho_0)dv-Q(0)=0.$$

By using this result and (I-3b) in the expression (I-3a), we find

$$\frac{\pi K}{\lambda^*}\int_{u_i}^{u_N}(\rho-\rho_0)dv=\int_{u_i}^{u_N}\frac{\delta du}{(u-u_i)^{\frac{1}{2}}}. \quad \text{(I-3d)}$$

To integrate the right-hand side by parts note that $d\delta/du$ exists and is continuous. This may be seen by integrating (I-2c) by parts and differentiating. The piecewise continuity of the derivatives of $(\rho-\rho_0)$ assures the existence and continuity of $d\delta/du$.

That $\delta(u_N, z)=0$ may be seen also from the form of (I-2c) after integration by parts. Therefore,

$$\frac{\pi K}{\lambda^*}\int_{u_i}^{u_N}(\rho-\rho_0)dv=-2\int_{u_i}^{u_N}(u-u_i)^{\frac{1}{2}}\frac{d\delta}{du}du;$$

differentiating with respect to u_i we obtain

$$(\rho-\rho_0)_{u=u_i}=-\frac{\lambda^*}{\pi K}\int_{u_i}^{u_N}\frac{(d\delta/du)du}{(u-u_i)^{\frac{1}{2}}}$$

which is the desired form.

5. Treatment of Discontinuities in $\rho-\rho_0$

The previous solution (I-3) can be extended by relatively simple means to apply to $\rho-\rho_0$ with finite discontinuities. This is the situation commonly encountered in disturbances bounded by a shock wave, e.g., airflow about a supersonic missile.

We assume (1) $\rho=\rho(u, z)$ is piecewise continuous (see Fig. 22) and has piecewise continuous first and second derivatives, (2) $\rho(u_s-, z)-\rho(u_s+, z)=\rho_2$ at $u=u_s$ with $0<u_s\leq u_n$, $\rho-\rho_0=0$ for $u\geq u_N$ and, finally, (3) Eq. (I-2c) holds.

Then defining the following reduced functions.

$$\begin{aligned}(\rho-\rho_0)_{\text{red}}&=(\rho-\rho_0)-\rho_2 && 0\leq u\leq u_s,\\ (\rho-\rho_0)_{\text{red}}&=(\rho-\rho_0) && u_s<u\leq u_N,\\ \delta(u)_{\text{red}}&=\delta(u)-\rho_2'(u_s-u)^{\frac{1}{2}} && 0\leq u\leq u_s,\\ \delta(u)_{\text{red}}&=\delta(u) && u_s<u\leq u_N,\end{aligned}$$

and

$$\rho_2'=\frac{2K}{\lambda^*}\rho_2;$$

we shall prove that

$$(\rho-\rho_0)_{i\text{red}}=\frac{\lambda^*}{\pi K}\int_{u_i}^{u_N}\frac{d\delta_{\text{red}}/du}{(u-u_i)^{\frac{1}{2}}}du. \quad \text{(I-4)}$$

By using (I-2c) and the definitions given above we obtain

$$\delta(u_i, z)=\frac{K}{\lambda^*}\left\{\int_{u_i}^{u_s}\frac{[(\rho-\rho_0)_{\text{red}}+\rho_2]dv}{(v-u_i)^{\frac{1}{2}}}+\int_{u_s}^{u_N}\frac{(\rho-\rho_0)_{\text{red}}dv}{(v-u_i)^{\frac{1}{2}}}\right\},$$

or upon integrating and rearranging

$$\delta(u_i, z)-\rho_2'(u_s-u_i)^{\frac{1}{2}}=\frac{K}{\lambda^*}\int_{u_i}^{u_N}\frac{(\rho-\rho_0)_{\text{red}}dv}{(v-u_i)^{\frac{1}{2}}}, \quad \text{(I-4a)}$$

$$\delta(u_i, z)_{\text{red}}=\frac{K}{\lambda^*}\int_{u_i}^{u_N}\frac{(\rho-\rho_0)_{\text{red}}dv}{(v-u_i)^{\frac{1}{2}}}. \quad \text{(I-4b)}$$

(I-4) follows at once since the reduced function $(\rho-\rho_0)_{\text{red}}$ is continuous and has piecewise continuous first and second derivatives. The result of Sec. IV applies.

6. Evaluation of ρ_2'

We examine (I-4a), set $u_i=u$ and take the derivative with respect to $(u_s-u)^{\frac{1}{2}}$,

$$\frac{d\delta}{d(u_s-u)^{\frac{1}{2}}}-\rho_2'=-\frac{2K}{\lambda^*}(u_s-u)^{\frac{1}{2}}\int_u^{u_N}\frac{1}{(v-u)^{\frac{1}{2}}}\frac{d(\rho-\rho_0)_{\text{red}}dv}{dv},$$

after first integrating by parts and then differentiating. In the limit as $u\to u_s$ the right-hand side vanishes. If $u_s=u_N$ the same limit is approached; for $[d(\rho-\rho_0)_{\text{red}}]/dv$ is piecewise continuous and may be replaced by its

maximum. We conclude

$$\rho_2' = \lim_{u \to u_s} \frac{d\delta}{d(u_s-u)^{\frac{1}{2}}}. \tag{I-5}$$

If $u_s = u_N$, $\delta(u_N) = 0$ and

$$\rho_2' = \lim_{u \to u_N} \frac{-\delta(u)}{(u_N-u)^{\frac{1}{2}}}.$$

Approximate values of ρ_2' may be obtained from experimental data by plotting $-\delta(u)/(u_N-u)^{\frac{1}{2}}$ *versus* u and finding the limit of this ratio graphically as $u \to u_N$.

7. Derivatives of the Fringe Shift $\delta(u, z)$

We evaluate here expressions for $d\delta/du$ and $d^2\delta/du^2$ to use in a later discussion of the errors involved in approximating the integral (I-3) or (I-4).

For $(\rho-\rho_0)_{\text{red}}$ we use (I-4b). Thus

$$\frac{d\delta_{\text{red}}}{du} = \frac{K}{\lambda^*} \int_u^{u_N} \frac{d(\rho-\rho_0)_{\text{red}}/dv}{(v-u)^{\frac{1}{2}}} dv. \tag{I-6}$$

$$\frac{d^2\delta_{\text{red}}}{du^2} = \frac{K}{\lambda^*}\left[-\frac{1}{(u_N-u)^{\frac{1}{2}}}\left(\frac{d(\rho-\rho_0)_{\text{red}}}{du}\right)_{u_N} + \int_u^{u_N} \frac{1}{(v-u)^{\frac{1}{2}}} \frac{d^2(\rho-\rho_0)_{\text{red}}dv}{dv^2}\right]. \tag{I-7}$$

The integrand in (I-7) has a singularity at $v=u$. Since $d^2(\rho-\rho_0)_{\text{red}}/dv^2$ is bounded the integral exists. We note also that (I-7) has a singularity at $u=u_N$.

From the definition of $\delta(u)_{\text{red}}$ it is plain that

$$\frac{d\delta}{du} = -\frac{\rho_2'}{2(u_s-u)^{\frac{1}{2}}} + \frac{d\delta_{\text{red}}}{du}, \tag{I-8}$$

and

$$\frac{d^2\delta}{du^2} = -\frac{\rho_2'}{4(u_s-u)^{\frac{1}{2}}} + \frac{d^2\delta_{\text{red}}}{du^2}. \tag{I-9}$$

8. Error in Approximation to the Integral

The integral (I-3) or (I-4)§ is evaluated numerically by subdividing the interval (r_i, r_n) into a suitable number of zones and approximating the derivative $d\delta/dr^2$ in each small zone. Consider the corresponding subdivision in $u(u_i, \cdots u_j, u_{j+1} \cdots u_N)$ and use as approximation to the derivative

$$\left(\frac{d\delta}{du}\right)_a = \frac{\delta_{j+1}-\delta_j}{u_{j+1}-u_j}. \tag{I-10}$$

Somewhere in the interval (u_j, u_{j+1}), the derivative will take on the approximate value (I-10). Letting $u=a$ for this point we may obtain $d\delta/du$ by

$$\left(\frac{d\delta}{du}\right)_v = \left(\frac{d\delta}{du}\right)_a + \left(\frac{d^2\delta}{du^2}\right)_{u(\theta)} (v-a), \tag{I-10a}$$

where $u(\theta) = [a+\theta(u-a)]$ and $0 < \theta \leq 1$.

§ N.B. In this discussion we use δ without subscript but imply the continuity properties of δ_{red}. In actual calculations with the bound obtained, δ_{red} should be used.

The approximation to (I-3) may be written

$$(\rho_i/\rho_0 - 1)_a = -\Gamma \sum_{j=i}^{N-1} \int_{u_j}^{u_{j+1}} \left(\frac{d\delta}{du}\right)_a d(v-u_i)^{\frac{1}{2}}, \tag{I-11}$$

and the precise value given by

$$(\rho_i/\rho_0 - 1) = -\Gamma \sum_{j=i}^{N-1} \int_{u_j}^{u_{j+1}} \times\left[\left(\frac{d\delta}{du}\right)_a + \left(\frac{d^2\delta}{du^2}\right)_{u(\theta)} (v-a)\right] d(v-u_i)^{\frac{1}{2}} \tag{I-12}$$

if we use (I-10a) and for convenience set

$$\Gamma = 2\lambda^*/\pi K \rho_0.$$

Then the error in the approximation to the ratio ρ_i/ρ_0 is the difference of (I-12) and (I-11), thus

$$E = \left(\frac{\rho_i - \rho_{ia}}{\rho_0}\right) = -\Gamma \sum_{j=i}^{N-1} \int_{u_j}^{u_{j+1}} \times\left(\frac{d^2\delta}{du^2}\right)_{u(\theta)} (v-a) d(v-u_i)^{\frac{1}{2}}. \tag{I-13}$$

We note from (I-7) that $d^2\delta/du^2$ has a singularity at $u=u_N$. Multiplying the derivative by $(u_N-u)^{\frac{1}{2}}$ will remove this; so,

$$E \leq \Gamma \sum_{j=i}^{N-1} \int_{u_j}^{u_{j+1}} \times\left|\left(\frac{d^2\delta}{du^2}\right)_{u(\theta)} (u_N-v)^{\frac{1}{2}}\right| \frac{|v-a|}{(u_N-v)^{\frac{1}{2}}} d(v-u_i)^{\frac{1}{2}}, \tag{I-13a}$$

and

$$E \leq \frac{\Gamma}{2} \sum_{j=i}^{N-1} |\delta_{uu}(u_N-u)^{\frac{1}{2}}|_{\max} (u_{j+1}-u_j) \times\int_{u_j}^{u_{j+1}} \frac{dv}{(u_N-v)^{\frac{1}{2}}(v-u_i)^{\frac{1}{2}}} \tag{I-13b}$$

with the derivative written δ_{uu} for convenience. If we now regard each interval in r space as equal, we set $r_j = jw$, where w is the zone width and j the zone number. Thus,

$$u_{j+1} - u_j = w^2(2j+1) \leq w^2(2N-1) \leq w^2 2N. \tag{I-13c}$$

With inequality (I-13c)

$$E \leq \Gamma N w^2 |\delta_{uu}(u_N-u)^{\frac{1}{2}}|_{\max} \int_{u_i}^{u_N} \frac{dv}{(u_N-v)^{\frac{1}{2}}(v-u_i)^{\frac{1}{2}}}, \tag{I-14}$$

wherein the integral is readily recognizable by comparison with (I-3c). We have with $w = r_N/N$

$$E \leq \frac{\pi\Gamma |\delta_{uu}(u_N-u)^{\frac{1}{2}}|_{\max} r_N^2}{N}. \tag{I-15}$$

Elastic Constants of Beta-Brass Single Crystals*

ROBERT A. ARTMAN† AND DONALD O. THOMPSON‡
Department of Physics, The State University of Iowa, Iowa City, Iowa
(Received October 8, 1951)

Measurements of Young's modulus and the rigidity modulus for a set of beta-brass single crystals of 55 atomic percent copper and for single crystals of beta-brass near the zinc-rich phase boundary reveal a dependence of the elastic constants upon composition. Comparison is made with previous results, whereby it is found that the moduli of beta-brass for different compositions may be obtained by the use of a linear interpolation in terms of the atomic ratio of copper to zinc. Derived quantities are generally related in this manner also. The value of Σ_{111} (Poisson's ratio) is shown to be a positive quantity, in contrast to that derivable from previous results.

Results of bending-torsion measurements are presented. The agreement between measured values and those calculated from the principal elastic constants provides assurance for the consistency of the measurements.

INTRODUCTION

THE beta-phase of the copper zinc system is the one which results from approximately equi-atomic mixtures of the two elements. Its crystal structure is body-centered-cubic with the consequence that its elastic behavior is specified in terms of three constants.[1] At room temperature the structure possesses a high degree of order.

During the past decade there have been several studies made of the elastic properties of beta-brass.[2–5] Among this group of studies, however, the agreement in the values of the elastic coefficients has not been too good, and supposedly could be due to differences in composition. In only one of these investigations, that by Lazarus, was the same set of crystals used to obtain all three elastic coefficients, and the compositions of the finished crystals known. In other investigations, the compositions stated were those of the stock material. Inasmuch as Lazarus' results pertained to crystals near the zinc-rich edge of the phase and a composition dependence of the elastic coefficients was suspected, it was felt worthwhile to investigate the coefficients at the copper-rich side of the phase and even slightly into the adjacent alpha- plus beta-phase. In doing so, cognizance was to be taken of the fact that the crystals used in such an investigation should not differ in composition from each other.

The elastic coefficients are conveniently obtained from the results of two independent measurements; one, a determination of Young's modulus and the other, a determination of rigidity modulus of a suitable set of crystals of the same composition. It must be observed, however, that when a rigidity modulus measurement is attempted by observing the twist of a crystal rod under axial torque, one of two, or neither of the rigidity moduli may be obtained unless certain precautions are taken.[1] To assure the measurement of the free modulus, the apparatus must be so constructed that no external bending moments are applied to the crystal, i.e., the crystal must be free to bend, as it does in general under an axial torque. An experimental arrangement was therefore devised which provided assurance that this condition was met.

It seemed interesting moreover to add to the apparatus a means for measuring the bending which occurred as a result of applied axial torque and check this result with that predicted by theory.[6]

EXPERIMENTAL PROCEDURE

I. Preparation of Crystals

Single crystals of beta-brass reported upon in this and the following paper were grown by the Bridgman method. The stock material was a piece of rolled and annealed beta-brass given to this laboratory by the American Brass Company Geometrically, the crystals were cylindrical rods, 0.635 cm in diameter, and variable in length from 7 to 11 cm. The growth time of the crystals was carefully controlled, inasmuch as it was found that this factor was important in obtaining crystals of the same composition. After normal growth the crystals invariably differed in composition from the stock material in that they were more copper-rich. A few crystals near the zinc-rich phase boundary (less copper-rich than the stock) were prepared by placing calculated amounts of spectroscopically pure zinc and pieces of stock beta-brass in the mold. These components were then fused in the furnace. The resulting brass rod was then inverted in the mold and a crystal grown. This procedure was

* A preliminary report of these results appears in a Letter to the Editor, J. Appl. Phys. **22**, 358 (1951).

† The material on Young's moduli is based on the first sections of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate College of the State University of Iowa, June, 1951. Now at Evansville College, Evansville, Indiana.

‡ First Lieutenant, U. S. A. F.

[1] R. F. S. Hearmon, Revs. Modern Phys. **18**, 409 (1946).

[2] Wayne Webb, Phys. Rev. **55**, 297 (1939).

[3] J. S. Rinehart, Phys. Rev. **58**, 365 (1940); Phys. Rev. **59**, 308 (1941).

[4] W. A. Good, Phys. Rev. **60**, 605 (1941).

[5] D. A. Lazarus, Phys. Rev. **74**, 1726 (1948); Phys. Rev. **76**, 545 (1949).

[6] E. Goens, Ann. Physik. **15**, 455 (1932); Ann. Physik. **16**, 793 (1933).

followed as it seemed likely to help produce crystals homogeneous in composition.

Subsequent to growth the crystals were etched, quantitatively analyzed, and examined metallographically. The orientation of each crystal's symmetry axis with respect to the crystallographic axes was determined by means of back-reflection Laue diagrams. Generally, six such diagrams were used per crystal, taken mostly on the ends of the crystals and the results averaged. These various tests showed whether or not the crystals were single, gave their compositions and orientations, and indicated the amounts of alpha-brass precipitated into the beta-brass crystal. The alpha-phase shows up as bronze-colored needles or ferns in the yellow background of beta-brass.

II. Methods of Measurement

The Young's modulus measurements were made by the familiar composite piezoelectric oscillator method.[7]

The rigidity modulus measurements were made using a refined torsion lathe. A crystal specimen is mounted horizontally, held rigidly at one end by a collet. To the free end is attached a torsion wheel. Above and to one side of the torsion wheel a lever bar is supported by a knife edge bearing at its center, so that one end of the lever arm is at a transverse distance equal to the radius of the torsion wheel away from the axis of the crystal. A connecting cable is then placed between the torsion wheel and the end of the lever arm just above. Diametrically opposed to this cable is tied another connecting cable which supports one end of another lever arm directly below. The remaining ends of the two levers are also tied together by a connecting cable. Axial torque is then applied to the crystal by a weight suspended from the center of the lower lever arm, the weight consisting of a cylindrical vessel containing about 10 kg of lead floating in a tank of water. To increase the torque water is allowed to flow from the tank, thus decreasing the buoyant force.

The "twist" and "bend" measurements are made by means of optical levers, using a mirror arrangement on the crystal as shown in Fig. 1. Two rings are affixed to the crystal by means of set screws and to each ring a pair of mirrors is attached. The mirrors on each ring are located 90° apart. Relative twists or bends are then measured by observing the relative deflection of the images of a cross-hair from the appropriate pair of mirrors.

It should be pointed out that these observations give only the component of bending in a plane defined by the axis of the crystal and the axis of the small round rods by which the mirrors are held to the ring. It is necessary in this experiment, in order to obtain the other bending component, to rotate the ring on the crystal by 90° and repeat the observations. Incidentally, another measurement of the rigidity modulus is obtained in this procedure. Also, it is worthwhile noting that by proper orientation of the mirrors and images of the cross hairs upon them, the bending and torsion measurements are entirely unaffected by each other.

[7] W. T. Cooke and W. F. Brown, Phys. Rev. **50**, 1158, 1165 (1936).

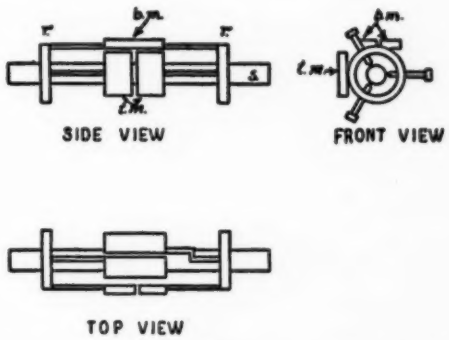


FIG. 1. Arrangement of mirrors on crystal specimen; b.m. mirrors used to observe bending; t.m. mirrors used to observe torsion.

According to theory[6] a crystal with a ⟨111⟩ orientation should show no bending when a pure axial torque is applied. A crystal having this orientation was investigated, with the result that absolutely no bending was observed. This negative result provides an adequate test of the apparatus, for if any bending had been observed with this specimen, it would have been due to an external bending moment produced by the apparatus. It may be noted that pieces of rolled stock brass rods showed slight bendings in this apparatus, indicating that they were not completely isotropic as might be supposed.

EXPERIMENTAL RESULTS

I. Young's Modulus and Rigidity Modulus

Young's modulus (E) is computed from experimentally obtained data by the equation

$$E=(2/n)^2\rho_{25}L_{25}^2(L_{25}/L_T)f_\beta^2[1+\tfrac{1}{2}An^2\pi\Sigma^2/L_{25}^2], \quad (1)$$

in which ρ_{25} and L_{25} are the density and length, respectively, of the beta-brass crystal at 25°C, L_T the length of the crystal at the temperature of measurement (presently $L_T=L_{25}$), Σ the average Poisson's ratio, n the number of half-waves in the crystal, A its cross-sectional area, and f_β the resonant frequency of the beta-brass crystal. In the case that L_T was different from L_{25} (following paper), L_T values were obtained from L_{25} values using the thermal expansion coefficients of polycrystalline beta-brass as determined by Merica and Schad.[8]

The free rigidity modulus (G) is computed from the data by the familiar equation for the isotropic modulus

$$1/G=\pi d^4\varphi/32lN, \quad (2)$$

in which d is the diameter of the specimen, l the length of crystal between the two mirror rings, φ the relative angular twist in radians, and N the applied axial torque.

[8] P. Merica and L. Schad, Bull. Bur. Standards **14**, 571 (1918).

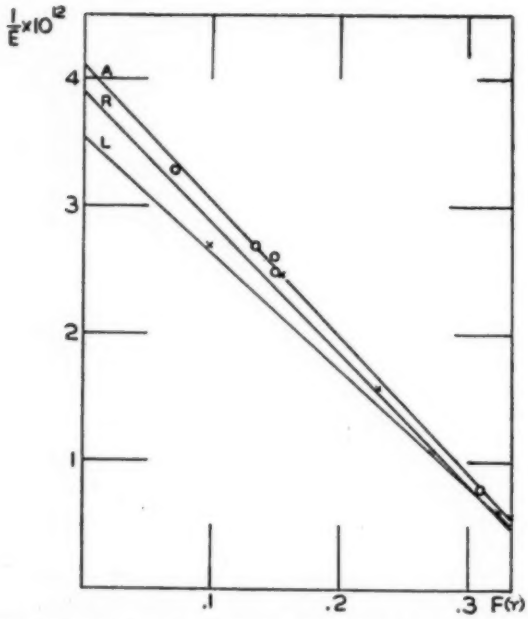


FIG. 2. The reciprocal of Young's modulus *vs* orientation function. Circles (○) designate crystals of nearly the same composition (55.12 atomic percent copper, set *A* of Table I). ×'s are for crystals of different composition. Curve *A* is drawn through the circle points, while curves *R* and *L* represent the previous data of Rinehart and Lazarus, respectively.

The principal elastic coefficients S_{11}, S_{12}, and S_{44} are related to Young's modulus and to the rigidity modulus

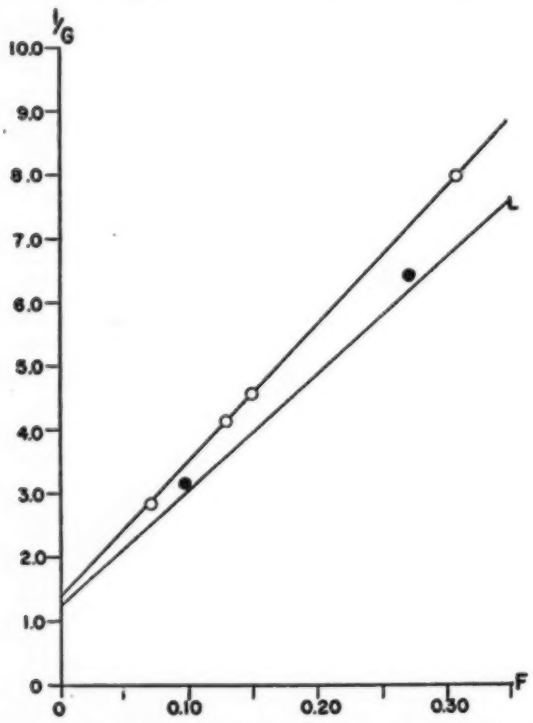


FIG. 3. The reciprocal of the rigidity modulus *vs* orientation function. Circles (○) designate crystals of nearly the same composition (55.12 atomic percent copper, set *A* of Table I). Filled circles (●) are for crystals of different composition. Curve *L* represents the data of Lazarus. The ordinate is $1/G\times10^{+12}$ cm²/dyne.

by the equations:

$$1/E=S_{11}-2SF(\gamma) \tag{3}$$

$$1/G=S_{44}+4SF(\gamma) \tag{4}$$

where

$$S\equiv S_{11}-S_{12}-\tfrac{1}{2}S_{44}. \tag{5}$$

The function $F(\gamma)$ is defined by

$$F(\gamma)\equiv\gamma_1^2\gamma_2^2+\gamma_1^2\gamma_3^2+\gamma_2^2\gamma_3^2, \tag{6}$$

in which the γ's are the direction cosines of the crystal's axis with respect to the crystallographic axes, and are obtained from the Laue diagrams.

The elastic parameters C_{11}, C_{12}, and C_{44} are related to

TABLE I. Orientation function, composition, reciprocal moduli, and measurements of the crystals.[a]

Crystal number	$F(\gamma)$	Atomic percent copper	$1/E$ cm²/dyne	$1/G$ cm²/dyne	Measurements and observations
7	0.310	55.50	0.785×10^{-12}	7.95×10^{-12}	*A*, *B*, *C*, (*P*)
8	0.133	54.96	2.690	4.13	*A*, *B*, *C*, (*P*)
9	0.154	54.45	2.454		*B*, *C*, *D*
10-*L*	0.149	55.86	2.497	4.625	*A*, *B*, *C*, (*P*)
10-*U*	0.149	55.23	2.610	4.48	*A*
12	0.325	56.72	0.6005 0.6222		*E*
14	0.070	54.85	3.287	2.85	*A*, *B*, *C*
15	0.230	54.12	1.572		*D*
16	0.098	51.88	2.687	3.17	*D*
17	0.333	70.25	0.571		*F*
18	0.272	52.57	1.090	6.41	*D*

[a] Index to Table I. Crystal No. 10 was long enough to permit two samples: 10-*L* is the lower section and 10-*U* is the upper section.
A Room temperature elastic coefficients.
B Room temperature internal friction.
C Temperature variation of Young's modulus and internal friction and the effect of precipitation of the alpha-phase on these quantities. Those crystals designated by (*P*) showed no alpha-phase after growth, but did show a small amount after several cycles of heating and cooling through the order-disorder transition.
D Composition dependence of the room temperature elastic coefficients.
E Young's modulus, internal friction, and their temperature variations for a crystal in the alpha- plus beta-phase. Initially the alpha-phase was not completely precipitated. $1/E$ was 0.6005 10^{-12} before heat treatment, and 0.6222×10^{-12} following heat treatment.
F Room temperature Young's modulus for an alpha-brass crystal.
Classifications *A*, *D*, and *F* are of interest in the present paper.
Classifications *B*, *C*, and *E* are of interest in the following paper (R. Artman, J. Appl. Phys. **23**, 475 (1952)).

the elastic coefficients in the following manner:

$$C_{11}=(S_{11}+S_{12})/(S_{11}-S_{12})(S_{11}+2S_{12}), \tag{7}$$

$$C_{12}=-S_{12}/(S_{11}-S_{12})(S_{11}+2S_{12}), \tag{8}$$

$$C_{44}=1/S_{44}. \tag{9}$$

It is clear from Eqs. (3) and (4) that plots of $1/E$ and $1/G$ *vs* $F(\gamma)$ yield the coefficients S_{11} and S_{44} as intercepts and have slopes proportional to the combination S. From this information S_{12} is calculated.

The orientation functions, compositions, $1/E$, and $1/G$ values for the various crystals are given in Table I. In accordance with the supposition that the elastic coefficients are composition dependent, a set of crystals of very nearly the same composition was chosen from all crystals grown. These are labeled *A* in Table I. The average composition of the set is 55.12 atomic percent

copper. The values of $1/E$ for these crystals are shown in Fig. 2 by circles and the $1/E$ values for crystals of a less copper-rich composition by x's. In Fig. 3 $1/G$ values for crystals of set A are shown by open circles, and $1/G$ values for two less copper-rich crystals by closed circles. Straight lines representing the $1/E$ values of Rinehart[3] and Lazarus[5] are included in Fig. 2, and the $1/G$ values of Lazarus[5] are included in Fig. 3. These lines are marked R and L, respectively.

The values of the elastic coefficients obtained from set A are given in Table II, along with the values resulting from previous investigations. S_{11} and S_{44} were obtained after fitting the $1/E$ and $1/G$ curves simultaneously to the same value of S by the method of least squares.

It should be noted that, although the Young's modulus measurements are adiabatic and the rigidity modulus measurements are isothermal in these experiments, no correction of the rigidity modulus is necessary. This is so because the adiabatic and isothermal rigidity moduli are the same.

TABLE II. Adiabatic elastic coefficients of beta-brass.[a]

	Rinehart	Lazarus[b]	Present	Ratio, present Lazarus
Atomic percent copper	52.47	51.74	55.12	[c]
$1/E_{100}=S_{11}$ (cm²/dyne)	3.881×10^{-12}	3.517×10^{-12}	4.105×10^{-12}	1.17
$1/G_{100}=S_{44}$		1.213	1.344	1.11
$1/E_{111}$	0.475	0.499	0.551	1.10
$1/G_{111}$		7.249	8.452	1.17
S	5.109	4.527	5.331	1.18
$-S_{12}$		1.616	1.898	1.18
			Average	1.15

[a] Good's values (see reference 4) are not included in the above table since it now seems certain that his value of S_{44} is too low.
[b] In transforming Lazarus' results to values of the S's (from C's) his directly measured value of $(C_{11}-C_{12})$, has been used rather than taking the difference in C_{11} and C_{12} as given in his paper.
[c] (55.12/44.88)/(51.74/48.26) =1.146, the ratio of relative atomic concentration.

II. Bending-Torsion Effect

If bending and torsion measurements are made over the same length of crystal, then an invariant elastic coefficient $\langle S_{34}'\rangle_{Av}$ in the laboratory coordinate system for a cylindrical specimen may be computed from the data by the equation:

$$\langle S_{34}'\rangle_{Av}=1/G[\theta_1^2/\varphi_1^2+\theta_2^2/\varphi_2^2]^{\frac{1}{2}}, \qquad (10)$$

where G is the measured rigidity modulus of the crystal, θ_1 the angular curvature of bending associated with the angular torsion φ_1, and θ_2, φ_2 the corresponding quantities measured in a plane perpendicular to that of θ_1 and φ_1.

According to crystal elasticity theory,

$$\langle S_{34}'\rangle_{Av}=2S[F(\gamma)-4F^2(\gamma)+3\gamma_1^2\gamma_2^2\gamma_3^2]^{\frac{1}{2}}, \qquad (11)$$

in which the quantities have all been defined previously. A comparison of values of $\langle S_{34}'\rangle_{Av}$ obtained from these two equations thus provides a good check as to the internal consistency of the measurements, for $\langle S_{34}'\rangle_{Av}$ as obtained from (10) depends upon the rigidity modulus

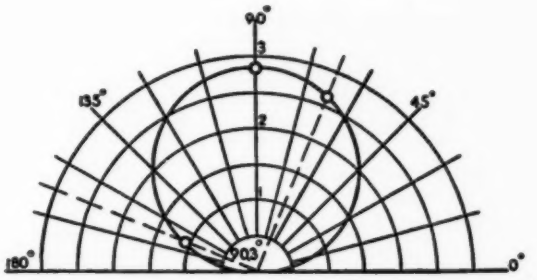


FIG. 4. Polar plot for crystal No. 10 of bending-torsion effect. Diameter of circle is the calculated value of $\langle S_{34}'\rangle_{Av}=2.91\times10^{-12}$ cm²/dyne.

result for the individual specimen and the bending-torsion angular ratios, whereas the value of $\langle S_{34}'\rangle_{Av}$ as obtained from (11) depends upon the slope value S of the whole set of crystals and the direction cosines found for the individual crystal.

Data obtained on crystals 8, 10 (average of both sections), and 14 are shown in Table III.

In Figs. 4 and 5, are the polar plots of these results. The diameter of the circle is the value of $\langle S_{34}'\rangle_{Av}$ from (11). Upon this circle are plotted two points whose radii vectores have magnitudes of $(1/G)(\theta_1/\varphi_1)$ and $(1/G)(\theta_2/\varphi_2)$. The check upon results is then imposed by measuring the angle between the radii vectores. This angle should be 90°. In all cases the measured angle lies well within the experimental error.

DISCUSSION OF RESULTS

The difference between the results of these experiments and those of Lazarus, is believed to be due entirely to a composition difference. Measurements given in Table I of other crystals which had compositions intermediate between the extreme concentration of Lazarus and the set A, show intermediate values of $1/E$ and $1/G$. It was suspected, therefore, that such values might agree with some form of linear interpolation since the concentration range is narrow. It will be noted in Table II that the ratios of the values of the coefficients (S_{ik}) obtained in the present experiments to those obtained by Lazarus are very similar, the average value being 1.15. This is to be compared to the value of the ratio of the relative atomic concentration, C_{Cu}/C_{Zn}, of the writers to the same quantity for Lazarus' alloy. It

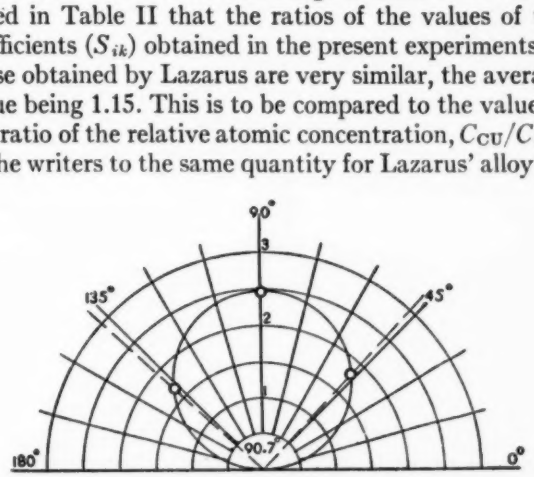


FIG. 5. Polar plot for crystal No. 14 of bending-torsion effect. Diameter of circle is the calculated value of $\langle S_{34}'\rangle_{Av}=2.48\times10^{-12}$ cm²/dyne.

TABLE III. Bending-torsion results.

Crystal No.	$\langle S_{44}' \rangle_{Av}$ (Eq. 10)	$\langle S_{44}' \rangle_{Av}$ (Eq. 11)
8	2.897×10^{-12} cm²/dyne	2.915×10^{-12} cm²/dyne
10	2.880	2.890
14	2.473	2.479

is 1.146 and, therefore, suggests the use of linear interpolation in terms of the atomic ratio of copper to zinc.

Table IV lists the relative atomic concentrations and reciprocal moduli of the crystals mentioned above. The values obtained by interpolation, marked computed, are in the main in satisfactory agreement with the experimentally observed values.

The fairly close agreement of ratios of reciprocal moduli to ratios of relative concentrations, as shown in Table IV, makes it appear as if, within the β-range, these various reciprocal moduli are directly proportional to the relative atomic concentration (C_{CU}/C_{ZN}). It is not, however, to be assumed that this relation should be used in an attempt to predict the behavior of β-brass if it could be maintained outside of its normal region. Almost any linear interpolation, such as one based on copper (or zinc) concentration alone will work about as well, simply because of the narrowness of the range, but would when extrapolated lead to very different predictions.

It will be noted that the values of $1/E$ obtained by Rinehart lie nearer the present data than to Lazarus' curve, even though the composition quoted by Rinehart is nearer to that of Lazarus. This is easily explained, however, for the compositions quoted by Rinehart are those of the stock material and not those of the crystals. A small loss of zinc during growth would completely account for the observed relation of his results to the other sets of data. Application of the above-named correction factor places the composition of Rinehart's crystals at 53.5 atomic-percent copper.[9]

TABLE IV. Comparison of computed and observed values for the moduli.

Crystal No.	$\frac{C_{CU}}{C_{ZN}}$	$1/E$ (comp.) cm²/dyne	$1/E$ (exp.) cm²/dyne	$1/G$ (comp.) cm²/dyne	$1/G$ (exp.) cm²/dyne
9	1.20	2.40×10^{-12}	2.45×10^{-12}		
15	1.18	1.60	1.57		
16	1.08	2.65	2.69	3.02×10^{-12}	3.17×10^{-12}
18	1.11	1.09	1.09	6.38	6.41

[9] A recent analysis by Dr. Rinehart of two of his crystals shows 53.4 atomic percent copper, J. Appl. Phys. **22**, 1089 (1951).

Table V lists interesting quantities which may be studied as a function of composition. In this table C' is the shear modulus $(C_{11}-C_{12})/2$, K is the bulk modulus, A is the anisotropy ratio defined as C_{44}/C', Σ_{100} is "Poisson's ratio" for a specimen subjected to a tensile stress along the $\langle 100 \rangle$ axis and Σ_{111} is the corresponding quantity for the $\langle 111 \rangle$ axis.

Perhaps the most significant quantity in Table V is C', a quantity whose value decreases with increasing copper concentration. This is consistent with theory, for as pointed out by Zener,[10] in a body-centered cubic copper lattice the contribution of the exchange interaction between conduction electrons and ions while positive is independent of the ion structure. Substitution of copper atoms for zinc would decrease the value of C'. The ratio of the two values of C' is nearly the same as the reciprocal of the previously mentioned interpolation ratio, 0.86 compared to 0.87. Furthermore, a smaller value of C' for copper-rich compositions means decreasing mechanical stability for the crystal structure as might be expected for compositions nearer the alpha-plus beta-phase boundary.§

TABLE V. Composition dependence of derived elastic constants.

	C' $\frac{\text{dynes}}{\text{cm}^2}\times10^{12}$	K $\frac{\text{dynes}}{\text{cm}^2}\times10^{12}$	A	Σ_{100}	Σ_{111}
Present	0.0833	1.079	8.933	0.462	0.220
Lazarus	0.0974	1.162	8.465	0.459	0.213

It is worth pointing out that either value of the anisotropy ratio A is considerably lower than that derivable from previous results,[4] and puts beta-brass in a class with Na and K in this respect. Also, either value of Σ_{111} is positive, whereas its value derivable from previous results[4] was negative and was the only exception in metals to positive values for this quantity.

The results of the bending-torsion measurements provide a test of internal consistency for the experiments, and confirm the previous statement that only a pure axial torque was applied to the crystals in measuring the rigidity moduli. If this were not the case, the agreement between the $\langle S_{44}' \rangle_{Av}$ values, if present at all, would be scattered.

The authors wish to express their indebtedness to Professor E. P. T. Tyndall for his suggestions and guidance throughout the course of the work.

[10] C. Zener, Phys. Rev. **71**, 846 (1947).

§ *Note added in proof:* In a Letter to the Editor to be published shortly, H. Jones shows that there are serious difficulties with the interpretation that the non-Coulombic ion-ion interaction which would result from the substitution of copper ions for zinc would cause the reduction of C'. The qualitative reason for the reduction of C' given in this paper should be disregarded.

Temperature Dependence of Young's Modulus and Internal Friction of Single Crystals of Beta-Brass*

ROBERT A. ARTMAN†
Department of Physics, The State University of Iowa, Iowa City, Iowa
(Received October 8, 1951)

The temperature variation of Young's modulus of beta-brass single crystals of 55 atomic percent copper has been determined. Comparison is made with previously reported results.

Internal friction has been measured at 21 and 42 kilocycles per second at room temperature and as a function of temperature to slightly above the complete disappearance of order. The room temperature decrement is for all crystals of the order of 10^{-4} and mainly independent of amplitude of oscillation. The decrement as a function of temperature shows a "relaxation peak." Zener's theory for alpha-brass is found to apply and the source of the relaxation is preferential orientation of pair axes of solute atoms in adjacent cells. The heat of activation is about 20×10^3 calories per mole. The relaxation peak is shown to be characteristic of the disordered state. There seems to be no relaxation phenomenon present in the ordered state.

The effect of precipitation of the alpha-phase on Young's modulus and the internal friction has been found negligible unless a large amount of precipitation occurs.

INTRODUCTION

THE present investigation started with several objects in view. The first of these consisted in determining the temperature variation of Young's modulus for a set of single crystals the composition of which was near the copper-rich edge of the phase boundary. This appeared worthwhile, in view of the variation of the room temperature elastic constants with composition reported in the previous paper. These results could then be compared with those reported by Rinehart.[1] The second object was to determine the temperature variation of the internal friction to and above the critical temperature of order-disorder. In addition, it was known that precipitation of inclusions of the alpha-phase might occur as a result of the heat treatment. The effect of the inclusions on Young's modulus and internal friction appeared to be a fruitful source of investigation.

EXPERIMENTAL PROCEDURE

The apparatus and methods of measurement were the same as those used by Swift and Richardson,[2] and Wert.[3] A quartz crystal is cemented to the brass specimen to form a composite piezoelectric oscillator. The method is such that Young's modulus may be obtained from the data necessary to determine the internal friction. The composite piezoelectric oscillator was supported on filaments of glass placed at displacement nodes, the quartz crystal being placed between two parallel brass plates which served as electrodes. These brass plates were mounted on a piece of transite, and the assembly was placed in the middle of a long glass tube. Electric connections were made by running leads through rubber stoppers in the ends of the tube. The side of the tube contained a lead to a vacuum pump, as all measurements were made at a pressure of 0.1 mm of Hg or less to prevent air damping. The assembly was heated by means of an electric furnace which surrounded the glass tube. Temperatures were measured with Chromel-Alumel thermocouples. Two materials were used to cement the quartz crystal to the brass specimen: phenyl salicylate for room temperature measurements and Insalute for temperature variation measurements.

One quartz crystal square in cross section with an area of 0.292 cm^2 was used in the composite piezoelectric oscillator throughout the course of the investigation. The crystal was used at both its fundamental frequency, 21 kilocycles per second, and its first harmonic. The measurement of the decrement‡ of a specimen depends on a prior knowledge of the electromechanical constant, K, and the decrement, Δ, of the quartz. The decrement of the quartz was determined by measuring the change in frequency, (Δf) when the bridge was balanced at a resistance twice that at resonance. The decrement is then given simply by

$$\Delta=2\pi(\Delta f)/f.$$

The electromechanical constant, K, in henries/gram can then be determined by

$$K=R/4\pi(\Delta f)m,$$

where R is the resistance of the composite oscillator at resonance, m = the mass of the quartz crystal, and (Δf) the change in frequency defined above.

Since it was desired to make measurements on the brass crystals at elevated temperatures, the quartz crystal had to be calibrated over the whole temperature range. This was done by making measurements of K and Δ both during the temperature rise and fall of the quartz between room temperature and 470°C. Measurements were made in approximately 25°C intervals, the

* Based on the last sections of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, in the Department of Physics in the Graduate College of the State University of Iowa, June, 1951.

† Now at Evansville College, Evansville, Indiana.

[1] J. S. Rinehart, Phys. Rev. **58**, 365 (1940); and **59**, 308 (1941).

[2] I. H. Swift and J. E. Richardson, J. Appl. Phys. **18**, 417 (1947).

[3] C. A. Wert, J. Appl. Phys. **20**, 29 (1949).

‡ The decrement, a measure of the internal friction, is the ratio of the energy dissipated per cycle to twice the total vibrational energy. It is related to Q by the relation $\Delta=\pi/Q$.

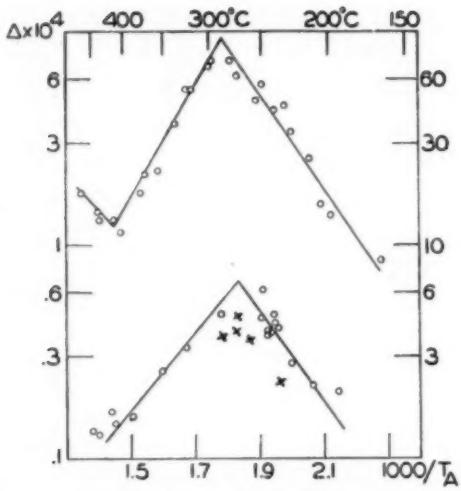


FIG. 1. Decrement of the quartz crystal *vs* temperature. The lower curve is the decrement of the quartz crystal at 42 kilocycles per second. The vertical scale is at the right of the figure. Data indicated by circles were obtained when the temperature was varied about 25°C over a period of several hours. Data indicated by crosses were obtained when the temperature was held constant for periods of time of the order of twelve hours. The upper curve is the decrement of the same crystal at 21 kilocycles per second. The vertical scale is on the left. Circles indicate data obtained by varying the temperature about 25°C over a period of several hours.

temperature changing through this increment over a period of several hours in most cases. The average value of three measurements taken at each temperature was used. The temperature gradient along the quartz increased from 0.5°C/cm at 200°C to 2°C/cm at 470°C. The rate of rise of temperature and gradient duplicated the same quantities in the runs taken with the composite oscillator. This was necessary for the calibration to be valid, as will appear later. The variation of the quartz decrement with temperature was particularly interesting and is described and discussed below.

EXPERIMENTAL RESULTS

Quartz

The results of the measurements of Δ and K are shown in Figs. 1 and 2, respectively. The lower curve is

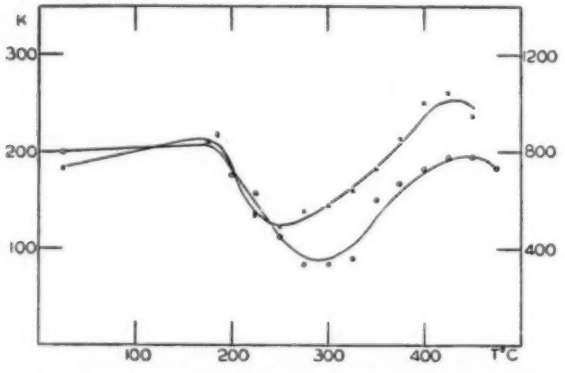


FIG. 2. Electromechanical constant K *vs* temperature. The circles (vertical scale at the left) represent data taken at 21 kilocycles per second. The X's (vertical scale at the right) represent data taken at 42 kilocycles per second.

the decrement of the quartz at 42 kilocycles per second. The vertical scale is at the right of the figure. Data indicated by circles were taken with the calibration procedure just described. Crosses indicate data taken with temperature kept constant for periods of 12 hours. The significance of the generally low values thus obtained will be discussed later. The upper curve is the decrement of the same crystal at 21 kilocycles per second. The vertical scale is on the left. Circles have the same significance as for the lower curve.

Young's Modulus§

Crystals classified as C in Table I in the preceding paper were used in the determination of the temperature variation of Young's modulus. A plot of observed

TABLE I. Temperature variation of Young's modulus in the principal directions.[a]

T(°C)	$1/E_{100}$	$1/E_{110}$	$1/E_{111}$
25	4.080	1.429	0.546
50	4.064	1.429	0.550
75	4.054	1.429	0.554
100	4.045	1.428	0.556
125	4.035	1.427	0.558
150	4.027	1.427	0.561
175	4.020	1.429	0.566
200	4.017	1.433	0.571
225	4.019	1.438	0.577
250	4.026	1.446	0.586
275	4.046	1.456	0.593
300	4.073	1.469	0.601
325	4.116	1.490	0.615
350	4.186	1.518	0.628
375	4.284	1.554	0.644
400	4.413	1.600	0.662
425	4.606	1.669	0.690
440	4.807	1.724	0.697
445	4.884	1.748	0.702
450	4.975	1.775	0.708
455	5.107	1.814	0.717
460	5.167	1.832	0.721
465	5.197	1.842	0.723
470	5.221	1.850	0.726
475	5.247	1.857	0.727
480	5.270	1.864	0.729

[a] The reciprocal moduli are $\times 10^{-12}$ cm²/dyne.

Young's modulus against temperature was made for each run for each of the crystals observed, and a smooth curve drawn through the points. Values of the modulus at 25°C intervals up to 425°C and in 5°C intervals from 440° to 480°C were read from these curves, approximately the same temperature intervals actually observed. Measurements were made at least twice on each crystal. From these values a line was fitted by the method of least squares at each temperature, and the values of $1/E_{100}$, $1/E_{110}$, and $1/E_{111}$ calculated. Plots of the relative variation of these results are shown in Fig. 3, and the values as calculated are shown in Table I. Relative variation is found by dividing the $1/E$ value at

§ The index of Table I of the preceding paper [Artman and Thompson, J. Appl. Phys. 23, 470 (1952)] applicable to this paper are as follows: B, C, E.

a particular temperature by the room temperature value.

It will be noted that a large change in the slope of the plots is observed at the critical temperature. The value of the critical temperature determined in this manner is 457°C. The results reported by Sykes and Wilkinson[4] for the transition temperature for the composition of this set of crystals is 457.5°C. The two sets of data appear to be in excellent agreement. Due to a difference in composition, Rinehart[1] observed a critical temperature of 468°C for his crystals. The results of the temperature variation of $1/E_{100}$, $1/E_{110}$, and $1/E_{111}$ are in essential agreement with those of Rinehart. His results are, however, smaller in magnitude than those reported here. Tabulation of Rinehart's room temperature and critical temperature values and their relative variation together with those obtained as a result of this work are given in Table II.

Rinehart's observations led him to conclude that there was no temperature hysteresis in Young's modulus of beta-brass. In the present work no temperature hysteresis was observed for three crystals, a slight amount for two others, and a permanent change in modulus after the first heating for crystal No. 12. This latter seems to be associated with the precipitation of alpha-brass in this crystal.

TABLE II. Relative variation of reciprocal moduli in the principal directions.[a]

	Rinehart[b]			Present		
Temp.	25°C	T_c	Relative variation	25°C	T_c	Relative variation
$1/E_{100}$	3.881	4.970	1.281	4.080	5.145	1.261
$1/E_{110}$	1.326	1.680	1.267	1.429	1.826	1.278
$1/E_{111}$	0.475	0.583	1.227	0.546	0.719	1.318

[a] The reciprocal moduli are $\times 10^{-12}$ cm²/dyne.
[b] See reference 1.

Internal Friction

In determining internal friction, plots were made of the temperature variation of the resistance of the composite oscillator at resonance. Smooth curves were drawn through the points, and resistance values read from these curves were used to obtain the internal friction. The decrement of four crystals (crystals No. 8, 9, 10-*L*, and 14) at 21 kilocycles per second are shown in Fig. 4, and for two others (crystals No. 7 and 12) at 42 kilocycles per second in Fig. 5.

Inspection of the results of the temperature variation of the decrement of beta-brass shows:

1. A pronounced peak at 315°C.
2. Relaxation strength, which is proportional to Δ_{max}, is independent of orientation function as shown by the close agreement of points for the four crystals in Fig. 4.
3. A lack of a discontinuity at the order-disorder critical temperature.
4. A shift of the peak to higher temperature for larger copper concentrations shown by comparing the results for crystals No. 7 and 12 in Fig. 5.
5. A lack of temperature hysteresis.

[4] C. Sykes and H. Wilkinson, J. Inst. Metals **61**, 223 (1937).

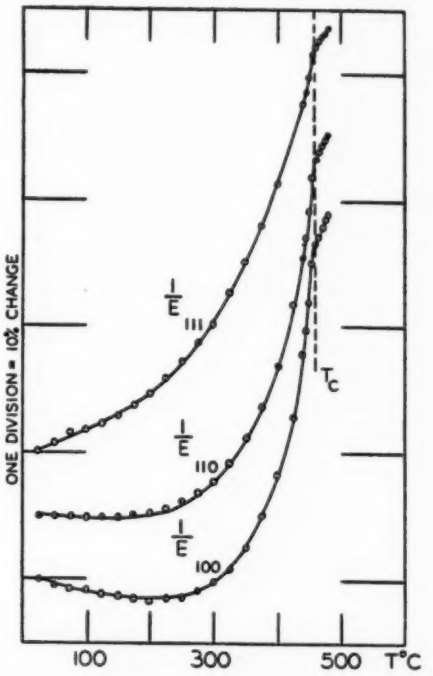


FIG. 3. Relative variation of reciprocal Young's moduli in the principal directions *vs* temperature. Each curve starts at a value of 1.000 at 25°C. The actual values of the reciprocal moduli are contained in Table I.

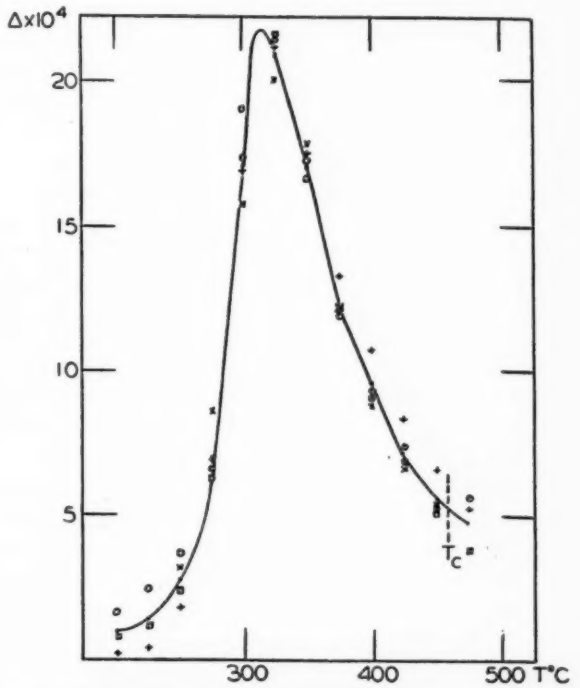


FIG. 4. Decrement of beta-brass at 21 kilocycles per second *vs* temperature. The symbols ○, □, +, and × represent data for crystals No. 8, 9, 10-*L*, and 14, respectively.

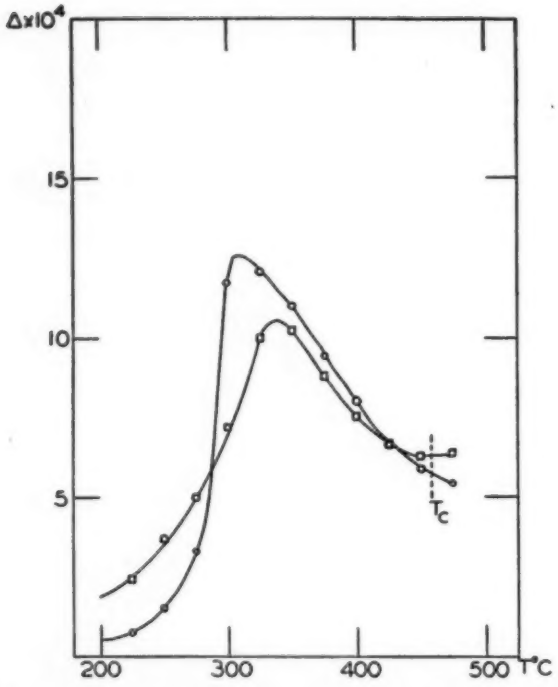


FIG. 5. Decrement of beta-brass at 42 kilocycles per second *vs* temperature. The symbols ○ and □ represent data for crystals No. 7 and 12, respectively.

The strain which may occur because of the differential expansion across the quartz-specimen interface is always a disturbing element in experiments of this kind. Effects of it seemed to appear here only on the first run of most of the crystals. The internal friction would increase rapidly and in a more or less erratic fashion from room temperature up to about 200°C. From this point on, however, consistent results were obtained. Furthermore, after the first run, the erratic behavior during the first 200°C was eliminated, and for temperature above 200°C, the measurements could be repeated upon subsequent runs to within five percent. Each of the specimens was vibrated in its fundamental mode. Since the crystals differed rather markedly in their lengths, the effect of a strain would have been more marked in the case of the shorter specimens. This was not observed. Moreover, heating the specimen would reduce any strains which might have been present. It is, therefore, concluded that the differential expansion at the quartz-specimen interface was not affecting the measurements above 200°C.

The decrement measurements were taken at as low a stress amplitude as was consistent with accurate observation. In all cases this was at a stress amplitude of the order of 1 to 2 g/mm². No attempt was made to make the measurements at constant stress amplitude. Instead, the amplitude dependence was checked at various points throughout a given run at stress amplitudes well above and below that used in the actual measurement with, in all cases, no change in the internal friction. This result was also obtained when decrement was systematically investigated as a function of stress amplitude at room temperature. The general effect was that the decrement curves were flat up to the largest stress amplitude used (in one case 380 g/mm²). This was true for all twelve cases taken at 21 kilocycles per second and for seven out of eleven cases at 42 kilocycles per second for the crystals designated by *B* in Table I of the preceding paper. No explanation is known as yet for the four exceptional cases of rising decrement at 42 kilocycles per second.

The temperature variation of the internal friction of beta-brass is significant when it is compared with the decrement measurements at room temperature. At a stress amplitude of 25 g/mm², the average value of 12 runs the decrement at room temperature for all the crystals was 5.9×10^{-5} at 21 kilocycles per second and 11.3×10^{-5} for 11 runs at 42 kilocycles per second. These values are approximately an order of magnitude less than the average results at 250°C. Furthermore, the decrement of the quartz crystal was 0.72×10^{-5} and showed no frequency effect and no stress amplitude effect when run alone. It appears that the frequency effect for the brass at room temperature though small is genuine.

Precipitation of the Alpha-Phase

Metallographic examination of the crystals after growth and before subsequent heat treatment showed no alpha-inclusions with the exceptions of crystals No. 12 and 17. Crystal No. 12 was in the alpha- plus beta-phase with the alpha-phase initially incompletely precipitated, while crystal No. 17 was completely in the alpha-phase. In each case where the temperature variation of Young's modulus and internal friction was determined, a section for metallographic comparison cut from the original crystal adjacent to the specimen section was placed in the furnace along with the specimen. After heat treatment comparison specimens from crystals No. 7, 8, and 10-*L* showed precipitation of a small

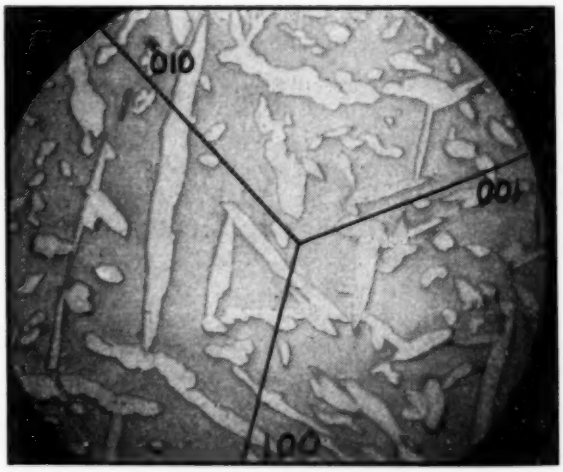


FIG. 6. Crystal No. 12 (×480) after heat treatment.

amount of alpha-brass, crystals No. 9 and 14 no precipitation, and crystal No. 12 complete precipitation of the alpha-phase.

A study was made by metallographic and x-ray methods, which need not be given in detail, of the shape and orientation of the precipitated alpha-brass in relation to the beta-matrix. When incompletely precipitated, the alpha-brass appeared as fine ferns, while complete precipitation resulted in structures similar to that shown in Fig. 6 (crystal No. 12 after heat treatment). The alpha-brass inclusions were found to be rods or needles with the long dimension usually parallel, or nearly so, to either the ⟨111⟩ or ⟨110⟩ direction of the beta-lattice. Mehl and Marzke[5] have previously reported the ⟨111⟩ direction.

Only in crystal No. 12, which was definitely within the alpha- plus beta-region, did the further precipitation of alpha-brass which occurred during the first temperature run produce any appreciable (>1 in 1000) change in the room temperature in Young's modulus. The value of $1/E$ increased by 3.6 percent. The change is in a direction which would occur if the alpha-orientation function were smaller than that of the beta-matrix. Other explanations are, however, possible.

The decrement also showed no essential change due to small precipitation, but again showed a large effect for crystal No. 12. The decrement declined from 60×10^{-5} (at 25 g/mm² stress amplitude) to 10×10^{-5} after the first temperature run. It seems likely that the decrease in decrement is due to the reduction of internal strains present in the crystal directly after growth and before the attainment of equilibrium in the structure.

DISCUSSION

A theory has been advanced by Zener[6] to explain the anomalous behavior of Young's modulus of beta-brass from −200°C to 200°C. The anomaly lies in $1/E_{100}$ and $1/E_{110}$ decreasing with temperature up to 200°C. Young's modulus in the three principal directions may be expressed as

$$1/E_{111}=1/9K+S_{44}/3,\quad 1/E_{110}=1/9K+S_{44}/4+1/12C',$$
$$1/E_{100}=1/9K+1/3C',$$

where K is the bulk modulus and C' is the shear modulus. It is noted that $1/E_{111}$ follows the normal behavior for its temperature variation, and from this is drawn the conclusion that K and S_{44} behave normally. Good's[7] data supports this conclusion. The anomalous behavior of $1/E_{100}$ is then attributed to the abnormal temperature variation of C'. Further, Zener[6] calculated an estimated temperature variation for C' based on the work of Fuchs, and it is found to be in good qualitative agreement with the temperature variation of $1/E_{100}$. Rinehart observed a decrease of 0.071×10^{-12} in $1/E_{100}$ from room temperature up to 200°C, whereas the decrease for the corresponding quantity in present data is 0.063×10^{-12}. At room temperature the contribution of $1/(3C')$ to $1/E_{100}$ is approximately forty times greater than that due to $1/(9K)$. The variation of $1/E_{100}$ with temperature is then, mainly, that due to C'. Comparison of the data on $1/E_{100}$ for the two compositions would seem to indicate that the change of $1/C'$ with temperature is smaller for higher copper concentrations.

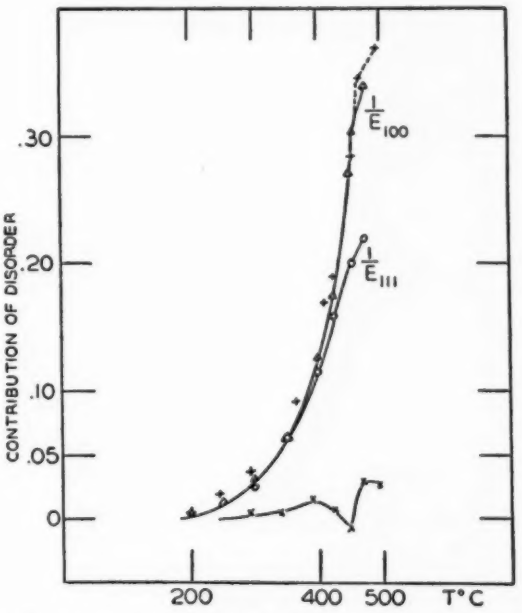


FIG. 7. "Contribution-of-disorder" *vs* temperature. The △'s and ○'s are data for $1/E_{100}$ and $1/E_{111}$, respectively, from present work, and the +'s and ×'s are data for the same quantities respectively from Rinehart's work.

Above 200°C a study of the temperature variation of reciprocal moduli is complicated by the onset of disordering, the variation indicating the effects of changing degree of order superimposed on the temperature variation of the ordered structure. One may, however, make comparisons of Rinehart's and present data above 200°C by noting that the relative variation of all principal reciprocal moduli are approximately linear from room temperatures up to 175°C. A straight line may be fitted to each of these curves over this temperature range. If one assumes that the straight line represents normal variation with temperature, then the difference of the values obtained from the fitted straight line and the actual relative variation curve represents the change in relative variation due to changing degree of order. For convenience, this difference will be called the contribution-of-disorder to the relative variations. Contribution-of-disorder curves were calculated for $1/E_{100}$ and $1/E_{111}$ for both Rinehart's and present data. The curves are shown in Fig. 7.

Comparison of the contribution-of-disorder curves for $1/E_{100}$ shows very close agreement between Rinehart's and present data, with Rinehart's curve being slightly

[5] R. F. Mehl and C. T. Marzke, Trans. Am. Inst. Mining Met. Engr. p. 123 (1931).
[6] C. Zener, Phys. Rev. **71**, 846 (1947).
[7] W. A. Good, Phys. Rev. **60**, 605 (1941).

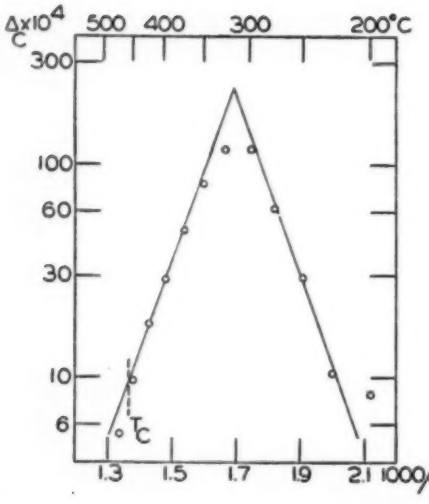


FIG. 8. "Corrected" values of the decrement of beta-brass at 21 kilocycles per second vs temperature. The points were calculated from the average curve in Fig. 4.

higher. At room temperature, the contribution of $1/(3C')$ to the value of $1/E_{100}$ is approximately forty times greater than that due to $1/(9K)$. If one assumes the ratio of these quantities remains approximately the same as the temperature is varied, then the contribution-of-disorder curve represents essentially the contribution-of-disorder of $1/(3C')$. The smaller change of $1/C'$ with higher copper concentrations for temperatures up to 200°C discussed earlier, leads to a smaller contribution-of-disorder curve at the higher temperatures. This may be noted in the curves in Fig. 7 where triangles ($\triangle$) represent data for present $1/E_{100}$ and plusses (+) that of Rinehart. The general conclusion can be drawn that the temperature variation of C' is only slightly sensitive to the copper concentration over the entire range of temperatures, provided the assumption that the ratio of K to C' does not change markedly with temperature is justified.

Comparison of the contribution-of-disorder curves for $1/E_{111}$ shows poor agreement between Rinehart and present data. The contribution of $S_{44}/3$ to $1/E_{111}$ is approximately four times as great as the contribution of $1/(9K)$ at room temperature. Assuming now that the ratio of S_{44} to K remains essentially unaltered as the temperature is varied, the $1/E_{111}$ contribution-of-disorder curve represents essentially the contribution of disorder of S_{44}. Comparison of the two curves would lead one to conclude that the temperature variation of S_{44} is highly sensitive to composition. However, Good[7] has measured the temperature variation of S_{44} directly. If one assumes that Good's measurements show the proper relative variation with temperature (even though the magnitudes may be incorrect), one may calculate the contribution-of-disorder of S_{44} for the relative variation curve. When this is done, it is found that this contribution-of-disorder curve follows closely, but is slightly higher than the $1/E_{111}$ curve of present data (represented by circles (O) in Fig. 7), and does not follow Rinehart's data in spite of the fact that both Good's and Rinehart's crystals had the same composition as indicated by critical temperatures. The most reasonable conclusion that can be drawn from comparisons of all data on $1/E_{111}$ is that the temperature variation of S_{44} is not highly sensitive to composition, and that Rinehart's variation of $1/E_{111}$ with temperature has an unexplained anomaly.

INTERNAL FRICTION

The results of the decrement of quartz as a function of temperature shown in Fig. 1 seem to support a theory outlined by Zener[8] in which the time of relaxation, τ, is assumed to vary with temperature through a heat of activation type of relation, i.e.,

$$\tau = \tau_0 \exp[H/(RT)].$$

The decrement as a function of temperature has then two limiting solutions, the one for large temperatures being proportional to $e^{H/RT}$ and that for small temperatures being proportional to $e^{-H/RT}$. Hence, if the logarithm of the decrement is plotted against $1/T$, the limiting solutions will give two straight lines whose slopes have a magnitude of H/R.

The measurements on quartz at both 21 kilocycles per second and 42 kilocycles per second fall reasonably well onto straight lines in the semilog plots. The slopes on the low temperature side of both the 21 and 42 kilocycles per second curves and on the high temperature side of the 21 kilocycles per second curve give a heat of activation of 10,400 cal/mole, while the slope of the high temperature side at 42 kilocycles per second gives a heat of activation of 8300 cal/mole. It appears that a reasonably accurate estimate of the heat of activation of the mechanism producing this phenomenon in the quartz may be placed at 10,000 cal/mole.

This relaxation peak, as mentioned above, did not appear stable‖ since the decrement declined when the temperature was held constant for periods of 12 hours as shown in Fig. 1 by the points plotted as x's. Since the data on quartz were obtained primarily as a calibration for the subsequent runs with quartz and brass together, the quartz data corresponding to the straight lines in Fig. 1 were used since these corresponded to the conditions under which the composite runs were taken.

The data for beta-brass shown in Figs. 4 and 5 indicate the presence of a relaxation peak. An attempt to calculate the heat of activation associated with the temperature dependence of the internal friction of beta-brass is complicated by the changing degree of order

[8] C. Zener, *Elasticity and Anelasticity of Metals* (University of Chicago Press, Chicago, Illinois, 1948).

‖ The nature of this peak is being investigated further at the State University of Iowa by Mr. George Alers. He has found that prolonged heating in a vacuum of the quartz crystal used by the writer will obliterate the peak completely. It seems that the peak may be due to absorption of water by the quartz. In this connection see Van Dyke and Palmer, Phys. Rev. 85, 745(A) (1952).

over the temperature range for which the peak occurs. A semilog plot of the data can be fitted fairly well with a straight line on either side of the peak. However, when this is done it is found that the slopes are quite different. Since this ignores the changing degree of order, a method was sought for injecting this fact into an interpretation of the data. The most successful method of accomplishing this was to assume that the decrement of the ordered phase was a slowly rising function of temperature without a peak as has been observed by Kê[9] for single crystals of pure aluminum, and that the peak was associated with the disordered phase. The assumption is then made that the internal friction of the disordered phase is proportional to the product of the measured decrement and the reciprocal of the degree of disorder. This has the effect of saying that if, at a given temperature, the degree of disorder is 0.20 that the observed result shows the effect of having the specimen only $\frac{1}{5}$ disordered at this temperature, and that the measurement for complete disorder at this temperature would then be proportional to five times that observed. This also has the effect of ignoring the temperature dependence of the ordered phase, but as the degree or order decreases with rising temperature, this contribution would become less important at the higher temperatures, even though the decrement for a completely ordered specimen would be rising. For the sake of definiteness, values calculated on this assumption will be called "corrected" values.

The results of applying the above analysis to the data are shown in Figs. 8 and 9. Order was calculated as a function of temperature, using the equation proposed by Cowley.[10] The curve at 21 kilocycles per second (Fig. 8) represents the values calculated from the average curve of Fig. 4, while in Fig. 9 circles (○) represent the data for crystal No. 7 and squares (□) the data for crystal No. 12, both at 42 kilocycles per second.

The curve in Fig. 8 is a typical symmetrical curve found wherever a single relaxation phenomenon is being dealt with. Since points on the low temperature side depend on rather large corrections for disorder, the agreement of the experiment with the assumption that the relaxation phenomenon is characteristic of the disordered structure seems well substantiated. The heat of activation is 19,500 cal/mole from the slope of either straight line. The results in Fig. 9 are not as gratifying. The two crystals agree on the high temperature side, where the disorder correction is small, but the heat of activation from the high temperature slope is only 13,700 cal/mole instead of 19,500 previously found. Perhaps this difference is no more than is to be expected on account of the neglect of the contribution of the ordered phase and its temperature variation, both of which may be different at 21 and 42 kilocycles per second. The curve for crystal No. 7 is somewhat asymmetric and, probably by accident, the slope of the

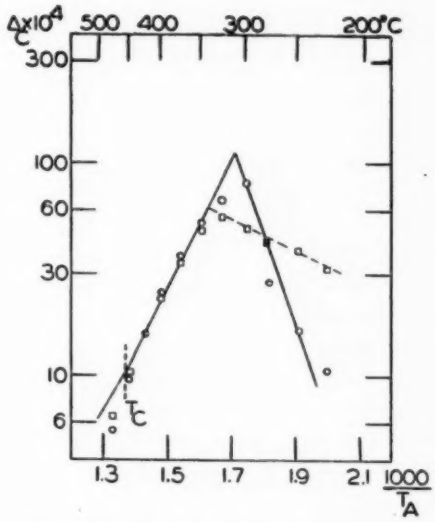


FIG. 9. "Corrected" values of the decrement of beta-brass at 42 kilocycles per second *vs* temperature. Circles (○) represent data for crystal No. 7, while squares (□) represent data for crystal No. 12.

low temperature side yields a value of H almost in exact agreement with the 19,500 cal/mol value obtained from Fig. 8. In crystal No. 12 the asymmetry is marked and the head of activation obtained from the low temperature side is only 3320 cal/mole. It must be remembered, however, that this crystal had considerable alpha-brass present which may complicate matters so much that the simple theory is inapplicable. It is quite possible that this variation is related to the precipitation of the alpha-brass. If order-disorder is the primary cause, it is rather difficult to explain why such complications are effective at lower temperatures and more order, and have little effect at higher temperatures and decreased order.

Zener[11] has developed a theory to explain a peak in the internal friction of an alpha-brass single crystal (of 30 percent zinc composition) which appears to be directly applicable to the disordered state of beta-brass. In this theory, the source of the internal friction is believed to be due to the relaxation of stress induced by preferential orientation of pairs of adjacent solute atoms. The beta-brass when disordered, may be thought of as a solution of copper in zinc (or zinc in copper) with, on the average, very nearly equal numbers of copper and zinc atoms per unit cell. In any individual cell, however, substitution of zinc for copper and copper for zinc will keep occuring on account of thermal agitation and diffusion. Replacing a zinc atom with a copper gives rise to an elastic distortion of the lattice, because copper atoms are smaller than zinc. If two adjacent cells contain substitution of copper for zinc, the distortion in the lattice will not have cubic symmetry. When a stress is applied to a crystal, the equilibrium distribution of the pair axes is not random. Thus, if a tensile stress is applied, the pairs will tend to be oriented preferentially along those permissible axes

[9] T. S. Kê, Phys. Rev. **71**, 533 (1947).
[10] J. M. Cowley, Phys. Rev. **77**, 669 (1950).
[11] C. Zener, Phys. Rev. **71**, 34 (1947).

which make the greatest angle with the tensile axis. This preferential distribution of pair axes will, however, require some time to be established, and consequently strain lags behind stress, resulting in internal friction.

In spite of the fact that there is a large difference in composition between Zener's alpha-brass crystal and the present crystals (30 atomic percent zinc compared to 45 atomic percent zinc), and that there is a change in structure from face-centered cubic to body-centered cubic, there are good reasons for believing that Zener's[11] theory is applicable to the present case. First, for the ordered state the only pair axes present would be those associated with the excess copper atoms. Since this number is small, the effect of pair axes in the ordered state will be small. Also, one would expect that the completely ordered structure will behave like a pure metal, and should show no peak in the internal friction *vs* temperature curve. Secondly, the "corrected" internal friction curves (Figs. 8 and 9) follow the theory well where the rate of change of order with temperature is large. It would be expected that if the theory were to break down, it would do so at the point where the correction factor undergoes a large rate of change with temperature. Instead, the agreement with theory seems best at this point. Thirdly, the relaxation strength of both the alpha-brass and the beta-brass are about the same order of magnitude. (Relaxation strength is defined as $2\Delta_{max}/\pi$.) Zener reports a relaxation strength of 0.024 for alpha-brass whereas present data give a relaxation strength of 0.0076 at 21 kilocycles per second, and 0.0044 for the average relaxation strength at 42 kilocycles per second. Finally, the heat of activation obtained (19,500 cal/mole) is consistent with the heat of activation for the diffusion of zinc in copper and for the heat of activation found by Zener[11] for an alpha-brass crystal, namely 33,000 cal/mole. Moreover, Rhines and Mehl[12] found a decrease from 40,000 to 31,000 cal/mole for the heat of activation of diffusion as the zinc concentration increased from 10 to greater than 16 percent.

A further clue to the sources of the internal friction may be obtained from the observation that relaxation strength is independent of orientation function. On the assumption that no relaxation will be induced by pure hydrostatic pressure, Zener[11] has developed the orientation dependence of relaxation strength. His result is

$$\Delta_E=\frac{(2/3)(S_{11}-S_{12})\delta'-2[(S_{11}-S_{12})\delta'-S_{44}\delta]F(\gamma)}{S_{11}-[2(S_{11}-S_{12})-S_{44}]F(\gamma)}$$

where

$$\delta=\frac{S_{44R}-S_{44U}}{S_{44U}} \quad \text{and} \quad \delta'=\frac{(S_{11}-S_{12})_R-(S_{11}-S_{12})_U}{(S_{11}-S_{12})_U}.$$

The subscript U refers to the unrelaxed and R to the relaxed parameter, and $F(\gamma)$ is the crystal orientation function. He further shows that in a body-centered cubic crystal δ' is identically zero if the source of the relaxation of δ' is nearest neighbors, i.e., atoms lying in the $\langle 111\rangle$ direction.

The observed results are that at 21 kilocycles per second the relaxation strength is independent of orientation from $F(\gamma)=0.07$ to $F(\gamma)=0.154$. If $\delta'=0$, calculation will show that the ratio of the relaxation strengths for $F(\gamma)=0.154$ to that for $F(\gamma)=0.07$ should be 2.85. Since this is well outside the experimental error, it is believed that $\delta'\neq 0$. Also, from an inspection of the equation giving the orientation dependence of relaxation strength it is obvious that if the relaxation strengths are equal for two orientations, they are equal for all orientations, and furthermore that $\delta'\neq 0$. Since experimentally the relaxation strengths at one frequency was found to be orientation independent in four cases, the evidence that $\delta'\neq 0$ is quite strong.

Since $\delta'\neq 0$, the relaxation source is not due to nearest neighbor exchanges. However, this does not exclude copper-zinc interchanges since next neighbors are not necessarily the other kind of atom. The reason that nearest neighbors do not exchange is probably that there is a good deal of short-range order preserved at and somewhat above the critical temperature. A zinc atom has a high probability of having copper atoms for its nearest neighbors (i.e., in the $\langle 111\rangle$ direction). To exchange a zinc atom for the copper atom destroys the short-range order and is unlikely to happen. The next to nearest neighbor relation contributes less to short-range order and this relationship need not be preserved. Consequently copper-copper and copper-zinc interchanges may continue to take place and pairs of neighboring distorted cells with pair axes may occur quite frequently. This then, is believed to be the source of the internal friction associated with the disordered phase of beta-brass.

In conclusion, the writer wishes to express his indebtedness to Professor E. P. T. Tyndall for his guidance and help during the progress of this work.

[12] F. N. Rhines and R. H. Mehl, Trans. Am. Inst. Mining Met. Engr. 128, 185 (1938).

Radiation Field of a Square, Helical Beam Antenna

H. LOTTRUP KNUDSEN
Microwave Laboratory, Royal Technical University of Denmark, Copenhagen, Denmark
(Received September 27, 1951)

Rigorous formulas have been derived for the field from a square, helical antenna with a uniformly progressing current wave of constant amplitude. These formulas that have the advantage of great simplicity are of direct use for helical antennas in the meter band, where for practical reasons only square helices are used. Further, in connection with corresponding rigorous formulas for the field from a circular, helical antenna with a uniformly progressing current wave of constant amplitude the present formulas may be used for an investigation of the magnitude of the error introduced in Kraus' approximate calculation of the field from a circular, helical antenna by replacing this antenna with an "equivalent" square helix. This investigation is carried out by means of a numerical example. The investigation shows that Kraus' approximate method of calculation yields results in fair agreement with the results obtained from the rigorous formulas. A statement to the contrary made recently in the literature is shown to rest on a trivial error.

INTRODUCTION

IN 1947 Kraus[1] described a helical antenna with such dimensions that the field radiated from this antenna has a sharp maximum in the direction of the axis of the helix and is circularly polarized in this direction. As yet no complete theory of the field radiated from this antenna has been given. However, by measurements Kraus[2,3] has found that the current distribution on a helical antenna with several turns with good approximation may be described as a uniformly progressing current wave with a constant amplitude. The problem of finding the field radiated by a helical antenna with this current distribution is mathematically well defined. Kraus[2,3] has given an approximate solution of this problem for a circular, helical antenna by first replacing the circular helix with an equivalent square helix and then making further mathematical approximations. The problem of the circular helix was solved rigorously by the author[4] and was later solved independently by Kornhauser.[5] In applying the rigorous formulas to a numerical example Kornhauser obtains a result which according to his statement differs considerably from the corresponding result obtained by using Kraus' approximate solution. However, it will be shown in this paper that there is a fair agreement between the results obtained by using Kraus' approximate formulas and by using the rigorous formulas. Kornhauser's erroneous statement rests on a trivial error that is pointed out in the following.

In the present paper rigorous expressions for the field from a square, helical antenna with a uniformly progressing current wave with constant amplitude will be derived. This work is essentially based on the last half of the author's aforementioned paper on circular and square helical antennas.[4] The obtained result will be of direct use for helical antennas in the meter band, helical antennas for this wavelength region being constructed as square helices. Further, in connection with the rigorous formulas for a circular, helical antenna, a knowledge of the rigorous formulas for a square, helical antenna will make possible an investigation of the magnitude of the error introduced in the calculation of the field from a circular helix by replacing this with an "equivalent" square helix as is done by Kraus[2,3] in his approximate method of calculation. This investigation will be carried out by applying the rigorous formulas for square and for circular helical antennas to a numerical example.

FORMULATION OF THE PROBLEM

Whereas a circular helix may be produced by folding a plane with a straight line on it into a circular cylinder, a "square" helix may be produced by folding a plane with a straight line on it onto a cylinder with a square cross section. In the present paper we set us the problem to calculate the field from a square helical antenna with a uniformly progressing current wave. For the sake of simplicity we confine ourselves to considering helical antennas with an integral number of turns. As shown by Kraus[2,3] a helical antenna with a progressing current wave may be considered an array of single turns. The array characteristic may be calculated by using well-known formulas; the problem therefore reduces to that of investigating the field from a single turn of the helical antenna.

A single turn of a square helix is shown in Fig. 1. A rectangular coordinate system (x, y, z) is introduced as shown in Fig. 1 and more clearly in Fig. 2, where the antenna is shown in two perpendicular projections. The antenna is made up of four pieces of straight wire connecting the points $(-b, 0, -2l)$, $(0, -b, -l)$, $(b, 0, 0)$, $(0, b, l)$, and $(-b, 0, 2l)$. The length of each of the straight pieces of line of which the antenna is composed is called c, the total length of a single turn L, and the

[1] J. D. Kraus, Electronics **20**, 109–111 (1947).
[2] J. D. Kraus and J. C. Williamson, J. Appl. Phys. **19**, 87–96 (1948).
[3] J. D. Kraus, Proc. Inst. Radio Engrs. **37**, 263–272 (1949).
[4] H. Lottrup Knudsen, *The Field Radiated by Circular and Square, Helical Beam Antennas* (Trans. Danish Acad. Tech. Sci., 1950), No. 8, 55 pp.
[5] E. T. Kornhauser, J. Appl. Phys. **22**, 887–891 (1951).

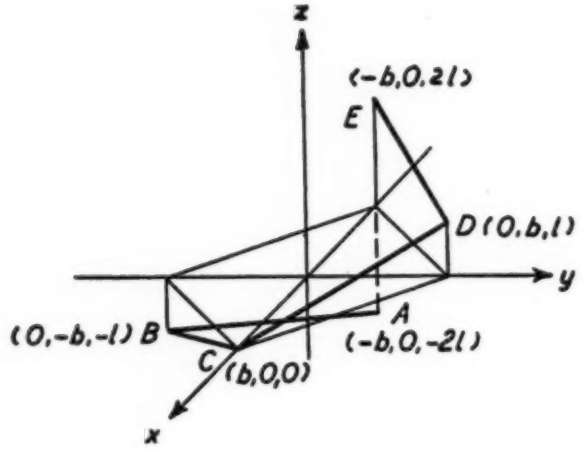


FIG. 1. Square, helical antenna with one turn.

pitch S. With these symbols we have

$$l/S=c/L=1/4, \tag{1}$$

$$c=(l^2+2b^2)^{\frac{1}{2}}. \tag{2}$$

Denoting the arc length of the antenna measured from the midpoint, C, of the wire and positive for $z>0$, by s, the current distribution on the antenna may be expressed by

$$I=I_0e^{i\beta s}; \tag{3}$$*

the parameter β denotes the transmission coefficient for the current wave. Introducing a terminology used by Kraus[2,3] we set $\beta=k/p$, where k is the intrinsic transmission coefficient of the medium surrounding the antenna.

CALCULATION OF THE RADIATION FIELD

We consider a current distribution in space with the density $J(\bar{r})$. Supposing that the current density differs from zero only in a finite domain we introduce a rectangular coordinate system (x, y, z) with its origo in or near this domain and a spherical coordinate system (r, θ, φ) oriented in the usual way with respect to the first-mentioned system. Letting $\zeta=\eta^{-1}$ denote the intrinsic impedance of space, r the distance from origo 0 to a point P in the far zone, $\bar{\rho}$ a unit vector pointing in the direction from 0 to P, and θ and φ the pole distance and the azimuth of P, we may express the electric field strength $\bar{E}$ and the magnetic field strength $\bar{H}$ at the point P in the following way:[4,6]

$$\bar{E}(r, \theta, \varphi)=T(r)\bar{F}(\theta, \varphi), \tag{4}$$

$$\bar{H}(r, \theta, \varphi)=\eta\bar{\rho}\times\bar{E}(r, \theta, \varphi), \tag{5}$$

with

$$T=Ki\omega e^{ikr}/r. \tag{6}$$

In these equations K denotes an arbitrary constant of which we shall despose later, whereas the spherical coordinates of the vector $\bar{F}$ are expressed by

$$F_r=0, \tag{7a}$$

$$F_\theta=F_x\cos\theta\cos\varphi+F_y\cos\theta\sin\varphi-F_z\sin\theta, \tag{7b}$$

$$F_\varphi=-F_x\sin\varphi+F_y\cos\varphi. \tag{7c}$$

The rectangular coordinates of $\bar{F}$ are defined by

$$F_x(\theta, \varphi)=\frac{1}{K}\frac{\mu}{4\pi}\int J_x(\bar{r}')\exp(-ik\bar{r}'\cdot\bar{\rho})dv'$$

$$=\frac{1}{K}\frac{\mu}{4\pi}\int J_x(x, y, z)\exp[-ik(x\sin\theta\cos\varphi+y\sin\theta\sin\varphi+z\cos\theta)]dxdydz, \tag{8}$$

and similar equations for F_y and F_z. In the first one of the above integrals $\bar{r}'$ denotes the position vector of the point at which the current element $J_x(\bar{r}')dv'$ is situated. In the last integral the suffix has been dropped as no misunderstanding seems to be possible.

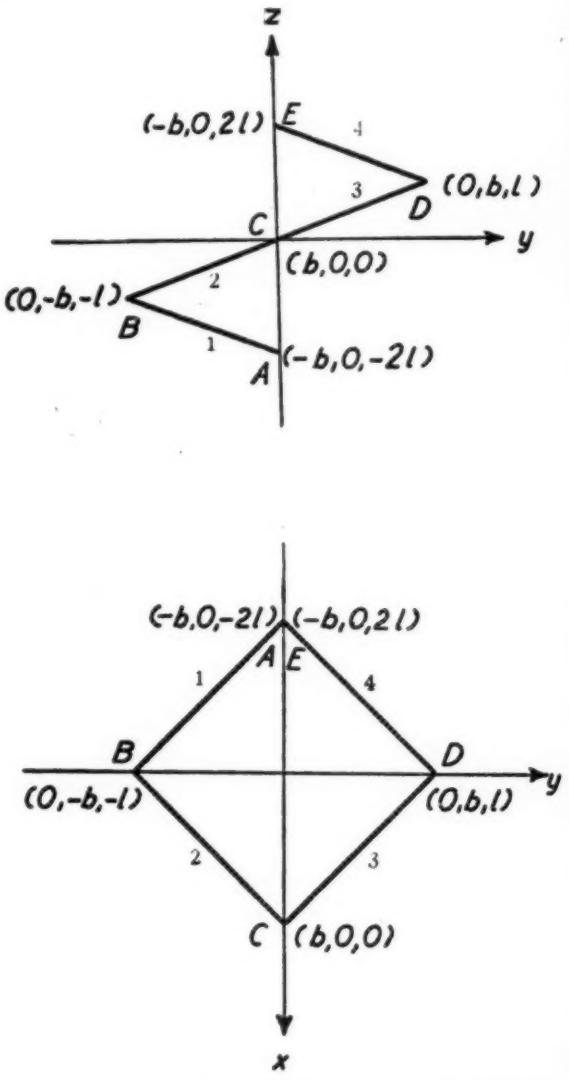


FIG. 2. Perpendicular projections of square, helical antenna with one turn.

* Time factor: $e^{-i\omega t}$.

[6] S. A. Schelkunoff, *Electromagnetic Waves* (D. van Nostrand Company, Inc., New York, 1943), pp. 331–336.

It follows from (4) that the electric field strength $\bar{E}$ may be obtained from the vector $\bar{F}$, by multiplying $\bar{F}$ with a complex scalar T depending on r, but being independent of θ and φ. For this reason we shall call $\bar{F}$ the normalized, electric field strength and T the normalization constant. The problem of finding the far zone field from a given current distribution will be completely solved when we have found the normalization constant T and the normalized, electric field strength $\bar{F}$.

In the present case of a single turn of a square helical antenna with the dimensions shown in Figs. 1 and 2 we find it convenient to choose

$$K=\mu I_0 b/4\pi. \tag{9}$$

The normalization constant then becomes

$$T=\frac{ikb\zeta I_0 e^{ikr}}{4\pi r}. \tag{10}$$

Further, numbering the four straight pieces of wire, of which the antenna is composed, from 1 to 4 as shown in Fig. 2, we express the normalized, electric field strength as a sum of four parts

$$\bar{F}=\sum_{\mu=1}^{4}\bar{F}^{\mu}, \tag{11}$$

where the μth part arises from the current in the μth wire of the antenna. We proceed to calculate the contribution to the normalized field strength coming from the first wire.

The straight piece of wire, 1, may be expressed by the following parametric equations

$$\left.\begin{aligned} x&=-\frac{b}{2}+bt\\ y&=-\frac{b}{2}-bt\\ z&=-\frac{3l}{2}+lt \end{aligned}\right\}-\tfrac{1}{2}\leqq t\leqq\tfrac{1}{2}. \tag{12}$$

A line element $d\bar{s}$ of the antenna is then expressed by

$$d\bar{s}=\left[\bar{i}\frac{dx}{dt}+\bar{j}\frac{dy}{dt}+\bar{k}\frac{dz}{dt}\right]dt=[\bar{i}b-\bar{j}b+\bar{k}l]dt, \tag{13}$$

where $\bar{i}$, $\bar{j}$, and $\bar{k}$ are unit vectors pointing in the x, y, and z directions. The arc length s introduced above is expressed by the parameter t in the following way,

$$s=-\tfrac{3}{2}c+ct. \tag{14}$$

For the current in conductor 1 we therefore have the following expression:

$$I=I_0\exp\left[\bar{i}\frac{k}{p}(-\tfrac{3}{2}c+ct)\right]. \tag{15}$$

The components of $\bar{F}^1$ may now be calculated by using the general formulas for F_x, etc., previously given in (8). Introducing the value of K given in (9) we find

$$F_x{}^1=\int_{-\frac{1}{2}}^{\frac{1}{2}}\exp\left\{ik\left[\frac{1}{p}(-\tfrac{3}{2}c+ct)-\left(-\frac{b}{2}+bt\right)\sin\theta\cos\varphi-\left(-\frac{b}{2}-bt\right)\sin\theta\sin\varphi-(-\tfrac{3}{2}l+lt)\cos\theta\right]\right\}dt, \tag{16a}$$

$$F_y{}^1=-F_x{}^1, \tag{16b}$$

$$F_z{}^1=SF_x{}^1/4b. \tag{16c}$$

In analogy with what has been done in the case of a circular helix[4] we define the following secondary parameters:

$$z=kb\sin\theta, \tag{17}$$

$$H=\frac{k}{2\pi}\left[\frac{L}{p}-S\cos\theta\right]=\frac{1}{\lambda}\left[\frac{L}{p}-S\cos\theta\right]. \tag{18}$$

By using these parameters we may express the above integral in the following way:

$$F_x{}^1=\exp\left\{i\frac{1}{2}\left[-H\frac{3\pi}{2}+z(\cos\varphi+\sin\varphi)\right]\right\}\times\frac{\sin\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi+\sin\varphi)\right]}{\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi+\sin\varphi)\right]}. \tag{19}$$

The components of the normalized, electric field strength of the field emitted by each of the remaining three conductors may be calculated by using similar procedures. For conductor 2 we obtain $F_x{}^2=F_y{}^2=4bF_x{}^2/S$, where

$$F_x{}^2=\exp\left\{i\frac{1}{2}\left[-H\frac{\pi}{2}+z(-\cos\varphi+\sin\varphi)\right]\right\}\times\frac{\sin\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi-\sin\varphi)\right]}{\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi-\sin\varphi)\right]}. \tag{20}$$

For conductor 3 we find $F_x^3=-F_y^3=-4bF_x^3/S$, where

$$F_x^3=\exp\left\{i\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi-\sin\varphi)\right]\right\}\times\frac{\sin\frac{1}{2}\left[H\frac{\pi}{2}+z(\cos\varphi-\sin\varphi)\right]}{\frac{1}{2}\left[H\frac{\pi}{2}+z(\cos\varphi-\sin\varphi)\right]}. \quad (21)$$

For conductor 4 is obtained $F_x^4=F_y^4=-4bF_x^4/S$, where

$$F_x^4=\exp\left\{i\frac{1}{2}\left[H\frac{3\pi}{2}+z(\cos\varphi-\sin\varphi)\right]\right\}\times\frac{\sin\frac{1}{2}\left[H\frac{\pi}{2}+z(\cos\varphi+\sin\varphi)\right]}{\frac{1}{2}\left[H\frac{\pi}{2}+z(\cos\varphi+\sin\varphi)\right]}. \quad (22)$$

Expressions for the vectors $\bar{F}^\mu$ having been obtained, the normalized, electric field strength $\bar{F}$ is calculated from (11). It is convenient to introduce the parameters $v_1, \cdots, v_6$ through the equations

$$v_1=\frac{1}{2}\left[H\frac{\pi}{2}+z(\cos\varphi+\sin\varphi)\right], \quad (23)$$

$$v_2=\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi+\sin\varphi)\right], \quad (24)$$

$$v_3=\frac{1}{2}\left[H\frac{\pi}{2}+z(\cos\varphi-\sin\varphi)\right], \quad (25)$$

$$v_4=\frac{1}{2}\left[H\frac{\pi}{2}+z(-\cos\varphi-\sin\varphi)\right], \quad (26)$$

$$v_5=\frac{1}{2}\left[H\frac{3\pi}{2}+z(-\cos\varphi-\sin\varphi)\right], \quad (27)$$

$$v_6=\frac{1}{2}\left[H\frac{3\pi}{2}+z(\cos\varphi-\sin\varphi)\right]. \quad (28)$$

By using these parameters we get

$$F_x(\theta,\varphi)=\cos v_5\frac{\sin v_2}{v_2}+\cos v_3\frac{\sin v_4}{v_4}-\cos v_4\frac{\sin v_3}{v_3}-\cos v_6\frac{\sin v_1}{v_1}+i\left[-\sin v_5\frac{\sin v_2}{v_2}-\sin v_3\frac{\sin v_4}{v_4}-\sin v_4\frac{\sin v_3}{v_3}-\sin v_6\frac{\sin v_1}{v_1}\right], \quad (29a)$$

$$F_y(\theta,\varphi)=-\cos v_5\frac{\sin v_2}{v_2}+\cos v_3\frac{\sin v_4}{v_4}+\cos v_4\frac{\sin v_3}{v_3}-\cos v_6\frac{\sin v_1}{v_1}+i\left[\sin v_5\frac{\sin v_2}{v_2}-\sin v_3\frac{\sin v_4}{v_4}+\sin v_4\frac{\sin v_3}{v_3}-\sin v_6\frac{\sin v_1}{v_1}\right], \quad (29b)$$

$$F_z(\theta,\varphi)=\frac{S}{4b}\left\{\cos v_5\frac{\sin v_2}{v_2}+\cos v_3\frac{\sin v_4}{v_4}+\cos v_4\frac{\sin v_3}{v_3}+\cos v_6\frac{\sin v_1}{v_1}+i\left[-\sin v_5\frac{\sin v_2}{v_2}-\sin v_3\frac{\sin v_4}{v_4}+\sin v_4\frac{\sin v_3}{v_3}+\sin v_6\frac{\sin v_1}{v_1}\right]\right\}. \quad (29c)$$

The components in the spherical coordinate system of the normalized, electric field strength $\bar{F}$ is obtained by substituting into (7) the above expressions for its components in the corresponding rectangular coordinate system.

SQUARE HELICAL ANTENNA WITH $H=1$

In the case of a single turn of a circular helical antenna with a uniformly progressing current wave the radiated field will be circularly polarized in the positive direction of the axis if the secondary parameter H,

$$H=1/\lambda[(L/p)-S\cos\theta], \quad (30)$$

where the symbols λ, p, L, and S have a meaning corresponding to the one attached to them in the case of the square helical antenna, is equal to 1 for $\theta=0$.[3] In the present section we shall investigate the polarization in the positive direction of the axis of the field radiated from a single turn of a square helical antenna with a progressing current wave and with $H=1$ for $\theta=0$. For $H=1$ and $\theta=0$ we have $z=0$, from which it follows that $v_1=v_2=v_3=v_4=\pi/2$ and $v_5=v_6=3\pi/4$. Inserting these values in (29) we hereby find for a square helix with a single turn,

$$F_x(0,\varphi)=-i8/\pi, \quad (31a)$$

$$F_y(0,\varphi)=8/\pi, \quad (31b)$$

$$F_z(0,\varphi)=0. \quad (31c)$$

Substituting in the expression (7b) and (7c) for the θ- and φ-components of the normalized electric field strength $\bar{F}$ we find

$$F_\theta(0,\varphi)=-i\frac{8}{\pi}e^{i\varphi}, \quad (32a)$$

$$F_\varphi(0,\varphi)=\frac{8}{\pi}e^{i\varphi}. \quad (32b)$$

The field in the positive direction of the axis for a square helix with a single turn and with $H=1$ for $\theta=0$ is thus seen to be circularly polarized. As in the case of a circular helix we conclude here that the field from a square helix with $H=1$ for $\theta=0$ and with an integral, but otherwise arbitrary number of turns, is circularly polarized in the direction $\theta=0$ and that the field in this direction from a square helix with $H=1$ for $\theta=0$ and with a large but not necessarily integral number of turns is approximately circularly polarized.[3,4]

SPECIAL CASES OF A SQUARE HELICAL ANTENNA

The square helical antenna treated above involves as special cases a rectilinear wire with a progressing current wave and a small square loop carrying a constant current. As a controlling measure and in order to demonstrate the use of the formulas for a square helical antenna of arbitrary dimensions as developed above, these formulas will be applied to the two special cases mentioned here.

Rectilinear Antenna

A rectilinear antenna with the length L placed so as to coincide with the part of the z axis between $z=-(L/2)$ and $z=(L/2)$ and carrying a current wave progressing in the positive z direction, $I=I_0\exp[i(kz/p)]$, may be obtained from a square helical antenna with one turn, with the length L and with $b=0$, where $2b$ denotes the diameter of the square cross section of the cylinder on which the helix is situated. With the notation introduced above we have for such a helical antenna that $b=0$, $S=L$, and $z=0$, from which it follows that $v_1=v_2=v_3=v_4=v=H(\pi/4)$ and $v_5=v_6=3v$. The components F_x and F_y give no contributions to the field components, as they are finite, and as they are to be multiplied by the normalization constant T containing b^1, which is zero. The component F_z contains b^{-1}, and therefore, when multiplied by T, gives a finite contribution to the field components. We find

$$F_z=\frac{L}{4b}\frac{2\sin v}{v}(\cos 3v+\cos v)=\frac{S}{b}\frac{\sin 4v}{4v}. \tag{33}$$

Substituting this expression in (7) we get

$$F_\theta=-L\sin 4v\sin\theta/b4v, \tag{34a}$$

$$F_\varphi=0, \tag{34b}$$

from which we obtain by using (4)

$$E_r=E_\varphi=0, \tag{35a}$$

$$E_\theta=-\frac{i\zeta I_0 e^{ikr}}{2\pi r}\frac{\sin(kS/2)[(1/p)-\cos\theta)]}{(1/p)-\cos\theta}. \tag{35b}$$

This is just the well-known expression for the field in the far zone from a vertical antenna with a progressing current wave.[7]

Small Square Loop with Constant Current

A square loop with the diameter $2b\ll\lambda$ placed in the (x, y) plane with its center at origo and carrying a constant current I_0 may be obtained from a square helical antenna with one turn, with the height S equal to zero, and with the transmission coefficient β equal to zero. For such a helical antenna we have, with the notation introduced above, that $S=0$, $z=kb\sin\theta\ll 1$, and $H=0$, from which it follows that $v_1=-v_4=-v_5=(z/2)\times(\cos\varphi+\sin\varphi)\ll 1$ and $v_2=-v_3=-v_6=(z/2)(-\cos\varphi+\sin\varphi)\ll 1$. From this we find the following expressions for the rectangular components of the normalized, electric field strength,

$$F_x=i2\sin v_1\sin v_2\left[\frac{1}{v_1}+\frac{1}{v_2}\right]$$

$$\cong i2[v_1+v_2]=iz2\sin\varphi, \tag{36a}$$

$$F_y=i2\sin v_1\sin v_2\left[\frac{1}{v_1}-\frac{1}{v_2}\right]$$

$$\cong i2[-v_1+v_2]=-iz2\cos\varphi, \tag{36b}$$

$$F_z=0. \tag{36c}$$

After substituting in Eq. (7) we find $F_\theta=0$ and

$$F_\varphi=-i2z=-i2kb\sin\theta. \tag{37}$$

Using (4) we finally obtain $E_r=E_\theta=0$ and

$$E_\varphi=\frac{k^2A\zeta I_0 e^{ikr}}{4\pi r}\sin\theta, \tag{38}$$

where $A=2b^2$ is the area of the square loop. This is the well-known expression for the far zone field radiated by a constant current in a plane loop with the area A and with a diameter that is small compared to the wavelength.[8]

SQUARE, NUMERICAL EXAMPLE OF A HELICAL ANTENNA

The application of the above derived formulas will be demonstrated by a numerical example. The field will be calculated for a single turn of a helix with the same data as one of the helices used by Kraus.[2,3] Using the above introduced notation we may describe this helix by the following data, denoting by N the number of turns: $S=15.025$ cm, $L=72.26$ cm, $\lambda=66.67$ cm, and $N=7$. From these values of S and L we find $b=12.456$ cm and $l=3.756$ cm. For the parameter p, defined as the

[7] J. D. Kraus, *Antennas* (McGraw-Hill Book Company, Inc., New York, 1950), pp. 148–153.
[8] See reference 7, pp. 155–157.

quotient between the specific transmission coefficient of the medium surrounding the antenna and the transmission coefficient of the current wave, Kraus[3] gives the following semi-empirical formula occurring in Hansen and Woodyard's[9] theory of a homogeneous, linear array

$$p=\frac{L/\lambda}{(S/\lambda)+1+(1/2N)}. \tag{39}$$

For a helical antenna with the above given data we find $p=0.8358$. From the primary constants of the helical antenna in question we calculate the following secondary constants $z=67°.259\sin\theta$ and $H=1.2968-0.2254\cos\theta$.

By using the above given parameters we first calculate the x, y, and z components of the normalized, electric field strength $\bar{F}$ and then the θ- and φ-components of $\bar{F}$. The normalization constant being given by (10), the electric field strength $\bar{E}$ is then completely determined. In the following section we shall compare the field radiated from the square helical antenna considered in this section with the field radiated from an equivalent circular helical antenna having the same length and the same pitch and also the same current distribution along the wire as the square helix in question. In order to be able to compare the fields radiated from the two antennas we shall use the same normalization constant in both cases, *viz.*,

$$T=ika\zeta I_0 e^{ikr}/4\pi r, \tag{40}$$

where a denotes the radius of the cylinder on which the circular helix is situated. The normalized, electric field strength $\bar{F}^*$ corresponding to the normalization constant (40) is determined by $\bar{F}^*=b\bar{F}/a$. Through a simple calculation we find $b/a=1.1107$.

Calculations of F_θ^* and F_φ^* are made for $\varphi=0°$, 90°, 180°, and 270°. For each of these cases the real and imaginary parts of F_θ^* and F_φ^* are calculated and plotted for $0°<\theta<180°$. In Fig. 3† are plotted the real and the imaginary parts of F_θ^* as functions of θ for $\varphi=0°$ and $\varphi=180°$. In Fig. 4 are plotted the same functions for $\varphi=90°$ and $\varphi=270°$; for these values of φ the imaginary part of F_θ^* is zero. The real and the imaginary parts of F_φ^* are plotted in Fig. 5 as functions of θ for $\varphi=0°$ and $\varphi=180°$. In Fig. 6 are plotted the same functions for $\varphi=90°$ and $\varphi=270°$; for these values of φ the real part of F_φ^* is zero.

Kraus has used a simplified theory to calculate the field from a single turn of a helical antenna[3] which has—as far as the author can find out—the same data as the helix in the present numerical example, except that Kraus' antenna is a left-handed screw, whereas the present one is right-handed. In the notation of the present paper, Kraus has calculated and plotted the field components E_θ and E_φ‡ for the azimuthal angles $\varphi=90°$ and $\varphi=270°$. The reflected image in the axis $\theta=0$ of the curves obtained by Kraus are shown in Figs. 7 and 8. The reflection takes care of the opposite winding directions of the two helices. Except for the scale factor the curves in Figs. 4 and 7 showing the θ-component of the field should be identical, and so should the curves in Figs. 6 and 8 showing the φ-component. As may be seen from the figures there is not only a qualitative, but even a fair quantitative agreement between Kraus' approximate expression for the field from a square helix with a uniformly progressing current wave and the exact expression for the field given in this paper. No great error seems therefore to be committed in applying Kraus' approximate method of calculation for a rough estimation of the field from a single turn of a square helix. However, besides being more exact the present method has the advantage of greater clearness and simplicity and should therefore be preferred even for rough calculations. As Kraus[3] points out, for a helical antenna with several turns the effect upon the total field of the field from a single turn is overshadowed by the influence of the array characteristic. However, the polarization of the field from a helical antenna is exactly the same as the polarization of the field from a single turn of the antenna. In investigations of helical antennas it is therefore important that reliable formulas for the field from a single turn are available.

Kraus' approximate method[2] of calculating the field from a circular helical antenna involves two steps: (1) the circular helix is replaced by an equivalent square helix, (2) various mathematical approximations are introduced in the calculation of the field from the square, helical antenna. The legitimacy of the last step was discussed above. In the next section we shall investigate whether the first step is justified or not, by plotting curves of the field components of a circular, helical antenna equivalent to the square, helical antenna considered in the above numerical example, and comparing the curves calculated in this way with the curves obtained for the square helix.

THE FIELD RADIATED BY AN EQUIVALENT CIRCULAR HELICAL ANTENNA

Using the definition of equivalence of a square and a circular helical antenna given by Kraus[2] we investigate in this section the field from a circular helical wire with the same length, the same pitch, and the same current distribution as the square helix in the example given above. For this purpose we use the exact formulas for

[9] W. W. Hansen and J. R. Woodyard, Proc. Inst. Radio Engrs. **26**, 333–345 (1938).

† In Fig. 15 of the author's 1950 paper (reference 4), which corresponds to Fig. 3 of this paper, "*Re*" and "*Im*" should be interchanged. This error was kindly pointed out to the author by Professor J. D. Kraus.

‡ In Fig. 9 of Kraus' 1949 paper (reference 3) it seems that the symbols $E_{\theta T}$ and $E_{\phi T}$ should be interchanged. The mistake probably has its origin in the circumstance that the curves in the aforementioned paper were calculated on the basis of the method of calculation given in Kraus' 1948 paper (reference 2), where the helix had an orientation different from the one occurring in the first-mentioned paper.

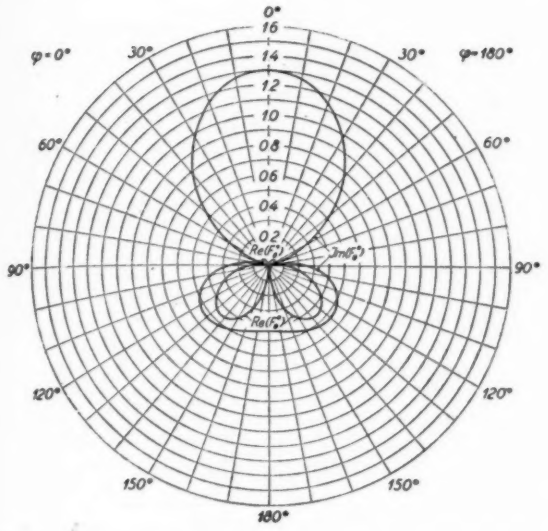


FIG. 3. The real and the imaginary part of the θ-component, F_θ^*, of the normalized, electric field strength of the field from the square helical antenna, plotted as functions of θ for $\varphi=0°$ and 180°. The curves were calculated according to the rigorous formulas of this paper.

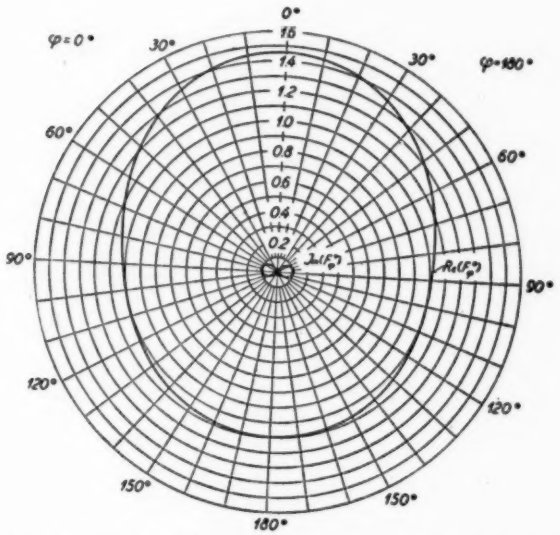


FIG. 5. The real and the imaginary part of the φ-component F_φ^* of the normalized, electric field strength of the field from the square helical antenna, plotted as functions of θ for $\varphi=0°$ and 180°. The curves were calculated according to the rigorous formulas of this paper.

the field radiated by a circular helix with a progressing current wave, developed by the author[4] and independently by Kornhauser.[5] With the normalization constant T given by (40) calculations of the θ- and φ-components of the normalized, electric field strength $\bar{F}$ are made for $\varphi=0°$, 90°, 180°, and 270°. For each of these cases the real and imaginary parts of F_θ and F_φ are calculated and plotted for $0°<\theta<180°$ in Figs. 9–12. There is a one to one correspondence between these four figures and the Figs. 3–6 in the mentioned order. A comparison between the plotted field components for the square and for the circular helical antenna show that there is a good qualitative and a fair quantitative agreement between the fields from the two equivalent antennas. It seems therefore that also the first step in Kraus' approximate calculation of the field radiated by a circular helical antenna is justified. However, as was pointed out by the author[4] and also by Kornhauser,[5] the exact formulas are clearer and simpler than Kraus' approximate formulas.

By considering the Figs. 9–12 we are now able to state the reason why Kornhauser[5] finds a considerable

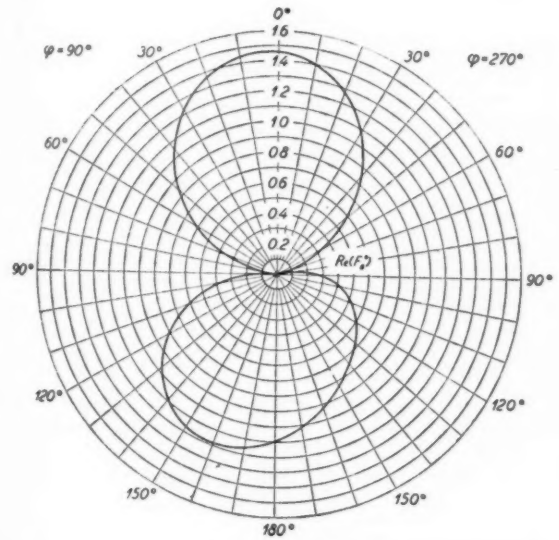


FIG. 4. The real part of the θ-component F_θ^* of the normalized, electric field strength of the field from the square helical antenna, plotted as a function of θ for $\varphi=90°$ and 270°; the imaginary part of F_θ^* is zero for these values of φ. The curve was calculated according to the rigorous formulas of this paper.

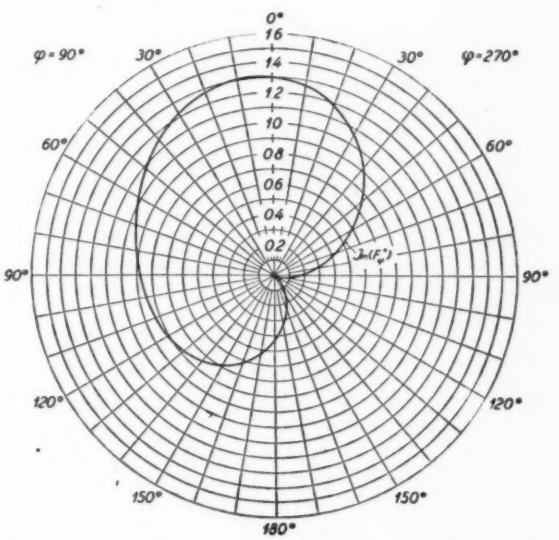


FIG. 6. The imaginary part of the φ-component, F_φ^* of the normalized, electric field strength of the field from the square helical antenna, plotted as a function of θ for $\varphi=90°$ and 270°; the real part of F_φ^* is zero for these values of φ. The curve was calculated according to the rigorous formulas of this paper.

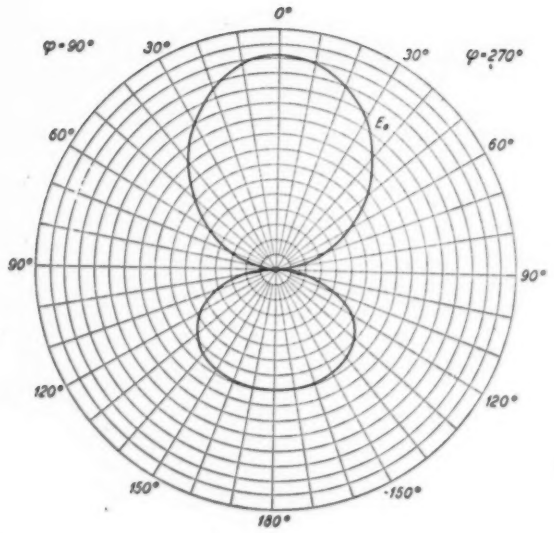


FIG. 7. The θ-component E_θ of the electric field strength of the field from the circular helical antenna, plotted as a function of θ for $\varphi=90°$ and 270°. The curve was calculated according to Kraus' approximate formulas.

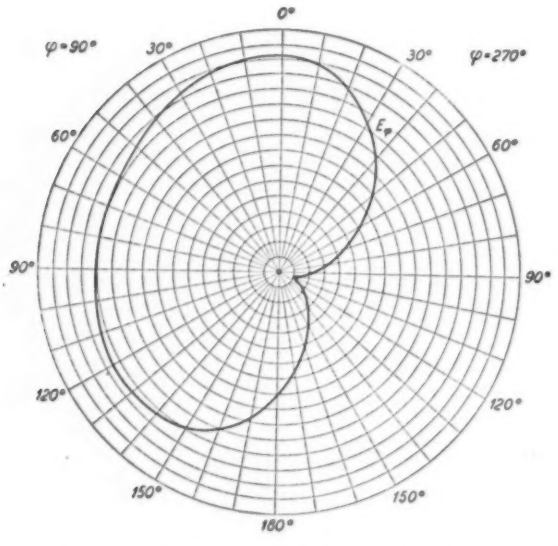


FIG. 8. The φ-component E_φ of the electric field strength of the field from the circular helical antenna plotted as a function of φ for $\varphi=90°$ and 270°. The curve was calculated according to Kraus' approximate formulas.

disagreement between the curves showing the field components calculated by using the exact formulas and the curves showing the field components calculated by using Kraus' approximate method. In the notation used in this paper Kornhauser compares in Fig. 4 of his paper the field components in the plane $\varphi=0°$ and $\varphi=180°$ calculated by the exact method with the field components in the plane $\varphi=90°$ and $\varphi=270°$ calculated by Kraus' method. The two sets of curves do not agree, and they should not.

CONCLUSION

For mechanical reasons a helical beam antenna for use in the meter-band is constructed as a square helix of wire. As shown by Kraus the current distribution on a helical antenna may approximately be described as a uniformly progressing current wave with a constant amplitude. In the present paper rigorous formulas are derived for the components of the field radiated by a square helical antenna with a current distribution of

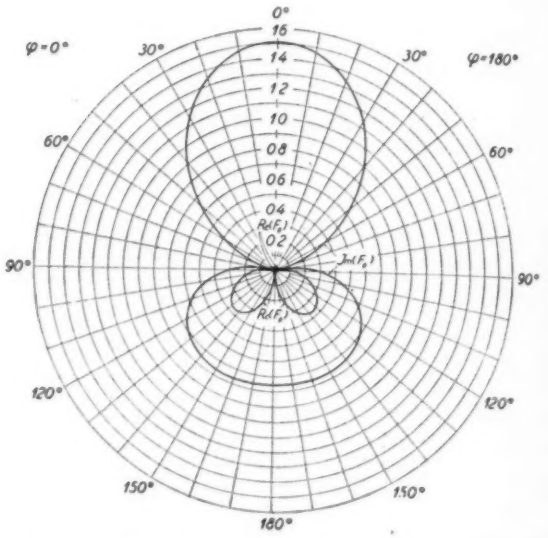


FIG. 9. The real and the imaginary part of the θ-component F_θ of the normalized, electric field strength of the field from the circular helical antenna, plotted as functions of θ for $\varphi=0°$ and 180°. The curves were calculated according to the rigorous formulas derived in a previous paper by the author.

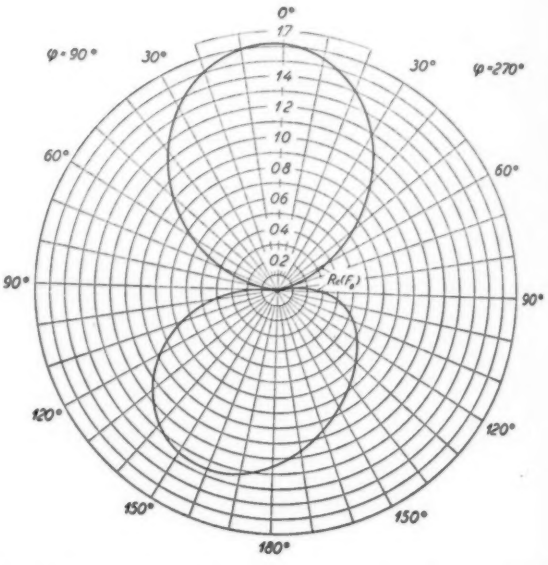


FIG. 10. The real part of the θ-component F_θ of the normalized, electric field strength of the field from the circular helical antenna, plotted as a function of θ for $\varphi=90°$ and 270°; the imaginary part of F_θ is zero for these values of φ. The curve was calculated according to the rigorous formulas derived in a previous paper by the author.

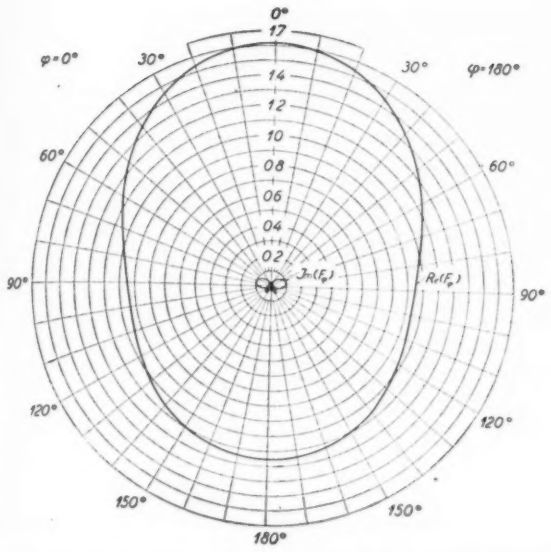


FIG. 11. The real and imaginary parts of the φ-component F_φ of the normalized, electric field strength of the field from the circular helical antenna, plotted as functions of θ for $\varphi=0°$ and 180°. The curves were calculated according to the rigorous formulas derived in a previous paper by the author.

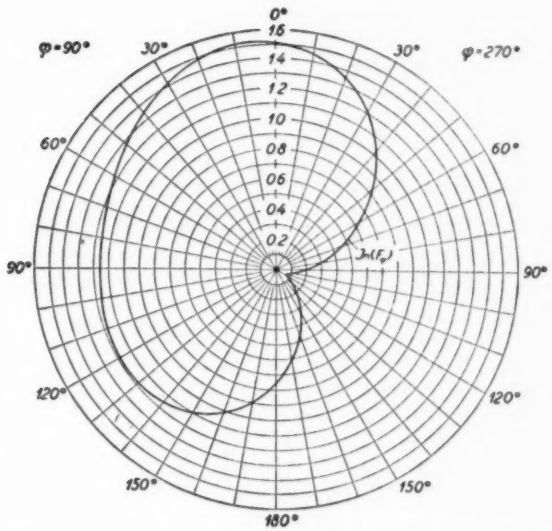


FIG. 12. The imaginary part of the φ-component F_φ of the normalized, electric field strength of the field from the circular helical antenna, plotted as a function of θ for $\varphi=90°$ and 270°; the real part of F_φ is zero for these values of φ. The curve was calculated according to the rigorous formulas derived in a previous paper by the author.

this type. These rigorous formulas turn out to be simpler than the approximate formulas derived by Kraus. Used on a numerical example the two sets of formulas show good agreement with each other.

In Kraus' approximate theory of the field from a circular helical antenna he first replaces the circular helix with an equivalent square helix having the same length, pitch, and current distribution. Using formulas derived in a previous paper by the author and independently by Kornhauser, the field radiated by a circular helical antenna equivalent to the square, helical antenna occurring in the above-mentioned, numerical example, was calculated. There is a good agreement, qualitatively as well as quantitatively, between the curves showing the components of the field from the equivalent, circular helix and from the original, square helical antenna. This points to the legitimacy of the first step in Kraus' approximate method of calculation.

By using the exact formulas for the field radiated by a circular helical antenna, Kornhauser comes to a result in apparent conflict with the corresponding result obtained by Kraus' approximate method. It is shown in the present paper that probably this disagreement is because of the fact that Kornhauser compares the field components in one plane, calculated according to the rigorous method, with the field components in a plane perpendicular to the first plane, calculated by using Kraus' approximate method.

ACKNOWLEDGMENT

The investigation presented here was carried out at the Microwave Laboratory of the Royal Technical University of Denmark, with a grant from the Research Council of the Technical Sciences in Denmark. I am indebted to the head of the Microwave Laboratory, Professor J. Oskar Nielsen for permission to publish this paper.

A Steady-State Heat Flow Problem for a Quarter Infinite Solid

WILLIAM KARUSH
The University of Chicago, Chicago, Illinois
(Received November 12, 1951)

The steady-state temperature field $T(x, y)$ in a homogeneous solid $x\geq 0$, $y\geq 0$, z arbitrary is studied when the flow out the face $y=0$ is a prescribed function $g(x)$. Along the face $x=0$ either the condition $T=0$ or $k(\partial T/\partial x)=hT$ is assumed. Of especial interest is the temperature at the face $y=0$ and the flow at the face $x=0$; particular attention is paid to the case $g(x)=0$ for $x>a>0$. The results have application to situations in which heat is being drawn off at the corner of a solid.

INTRODUCTION

THE purpose of this paper is to discuss a two-dimensional heat flow problem which arose in connection with certain applications but which may have some general interest. The problem concerns the steady-state harmonic temperature $T(x, y)$ in the quarter-infinite region $x\geq 0$, $y\geq 0$ resulting when a specified flow $g(x)$ of heat per unit length is drawn off (or put in) along the edge $y=0$ (see Fig. 1). Along the edge $x=0$ we suppose either (1) a constant temperature $T=0$ or, (2) a radiative flow hT with constant heat transfer coefficient h. Problems (1) and (2) will be treated separately although (1) may be regarded as the limit of (2) as $h\to\infty$. Our main concern is the character of the temperature along $y=0$ and the heat flow along $x=0$ with particular interest in flow distributions $g(x)$ which vanish for $x>a>0$. The above problem relates of course to a solid which is supposed to extend to infinity in both directions along the axis perpendicular to the xy-plane.

FIRST PROBLEM

Temperature 0 along One Face

In this problem we seek the solution $T(x, y)$ of the harmonic equation,

$$\nabla^2 T=0,$$

which satisfies the boundary conditions,

$$T=0 \quad \text{for} \quad x=0, \quad y\geq 0 \tag{1}$$

and

$$k\partial T/\partial y=g(x) \quad \text{for} \quad y=0, \quad x\geq 0. \tag{2}$$

Here k is the thermal conductivity of the solid. We wish a solution which tends to 0 for large y. A function which satisfies all our conditions but (2) is

$$T(x, y)=\int_0^\infty f(\alpha)\sin\alpha x e^{-\alpha y}d\alpha, \tag{3}$$

where f is arbitrary. To determine f we substitute into (2) obtaining

$$-k\int_0^\infty \alpha f(\alpha)\sin\alpha x d\alpha=g(x)$$

which by Fourier's integral theorem yields the solution

$$f(\alpha)=-\frac{2}{\pi k\alpha}\int_0^\infty g(t)\sin\alpha t dt. \tag{4}$$

We wish now to examine the temperature $T(x)$ along the edge $y=0$. To this end we substitute (4) into (3), set $y=0$, and interchange the order of integration. Using the standard result

$$\int_0^\infty \frac{\sin\alpha x\sin\alpha t}{\alpha}d\alpha=\tfrac{1}{2}\ln\left|\frac{t+x}{t-x}\right|,$$

we obtain

$$T(x)=-\frac{1}{\pi k}\int_0^\infty g(t)\ln\left|\frac{t+x}{t-x}\right|dt. \tag{5}$$

To proceed further we suppose that

$$g(x)=0 \quad \text{for} \quad x>a, \tag{6}$$

i.e., that the surface is insulated beyond $x=a$, and seek the average $\bar{T}$ of $T(x)$ over the interval $0\leq x\leq a$. From the known integral

$$\int_0^a \ln\left|\frac{x+t}{x-t}\right|dx=(a+t)\ln\left|\frac{a+t}{a-t}\right|+2t\ln\left|\frac{a-t}{t}\right|$$

we obtain

$$\bar{T}=-\frac{1}{\pi ak}\int_0^a g(t)[(a+t)\ln(a+t)$$
$$-(a-t)\ln(a-t)-2t\ln t]dt. \tag{7}$$

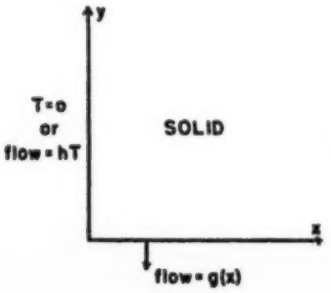


FIG. 1.

Next we consider the flow,

$$F(y)=-k\partial T/\partial x,$$

into the edge $x=0$, still assuming (6). From (3) and (4), and using the result

$$\int_0^\infty e^{-\alpha y}\sin\alpha t\,d\alpha=\frac{t}{y^2+t^2}$$

we obtain

$$F(y)=\frac{2}{\pi}\int_0^a \frac{tg(t)}{y^2+t^2}dt. \tag{8}$$

To obtain some information about the quantities (5), (7), and (8) we consider two special, simple forms for $g(x)$,

$$\begin{aligned}&\text{(a)}\quad g(x)=G, \quad 0\le x\le a,\\ &\text{(b)}\quad g(x)=Ax, \quad 0\le x\le a,\end{aligned} \tag{9}$$

where G and A are constants and (6) is assumed to hold. Denoting cases (a) and (b) by the subscripts C and L, respectively, we find

$$T_C(x)=-\frac{G}{k}\left[(a+x)\ln\frac{a+x}{|a-x|}+2x\ln\frac{|a-x|}{x}\right]$$

$$\bar T_C=-2\ln 2\frac{Ga}{\pi k}$$

$$F_C(y)=\frac{G}{\pi}\ln\left(\frac{a^2+y^2}{y^2}\right)$$

$$T_L(x)=-\frac{A}{\pi k}\left[\frac{a^2-x^2}{2}\ln\frac{a+x}{|a-x|}+ax\right]$$

$$\bar T_L=-\tfrac{2}{3}(\tfrac{1}{2}+\ln 2)\frac{Aa^2}{\pi k}$$

$$F_L(y)=\frac{2A}{\pi}\left[a-y\tan^{-1}\frac{a}{y}\right].$$

To see how sensitive $\bar T$ is to the exact form of $g(x)$ we compare $\bar T_C$ and $\bar T_L$ for flows in which the same total quantity of heat $\int_0^a g(x)dx$ is drawn off, i.e., distributions (a) and (b) for which $Aa=2G$. We find that $\bar T_L/\bar T_C=1.15$.

SECOND PROBLEM

Flow hT at One Face

Here we seek a harmonic function $T(x, y)$ which satisfies the conditions of the first problem except that the boundary condition (1) is replaced by

$$k\frac{\partial T}{\partial x}=hT \quad \text{for} \quad x=0, \quad y\ge 0. \tag{10}$$

We take the solution in the form,

$$T(x,y)=\int_0^\infty f(\alpha)(\sin\alpha x+\alpha p\cos\alpha x)e^{-\alpha y}d\alpha, \quad p=\frac{k}{h}, \tag{11}$$

this harmonic function satisfying (10). The unknown function f is again determined by (2), which leads to the integral equation.

$$-k\int_0^\infty \alpha f(\alpha)(\sin\alpha x+\alpha p\cos\alpha x)d\alpha=g(x). \tag{12}$$

This equation will be solved in the next section; for the moment we use the result obtained there, which yields, by Eq. (14), the solution

$$f(\alpha)=-\frac{2}{\pi k\alpha(1+p^2\alpha^2)}\int_0^\infty g(t)(\sin\alpha t+\alpha p\cos\alpha t)dt. \tag{13}$$

We again consider the two cases (9). It is possible to express the surface temperatures $T_C(x)$ and $T_L(x)$ at $y=0$ in terms of elementary and tabulated functions, but we omit the result because of its complexity. For the average temperatures we obtain

$$\bar T_C=-\frac{2Ga}{\pi k}\left\{\left(1+\frac{1}{q}\right)^2\ln 2+\frac{1}{q^2}[2e^q\mathcal{Ei}(-q)-e^{2q}\mathcal{Ei}(-2q)-\ln\gamma q]\right\},$$

$$\bar T_L=-\frac{2Aa^2}{\pi k}\left\{\left(\frac{1}{3}+\frac{1}{q}+\frac{1}{q^2}\right)\ln 2+\frac{1}{6}+\frac{1}{2q}+\frac{1}{q^2}[e^q\mathcal{Ei}(-q)-e^{2q}\mathcal{Ei}(-2q)]-\frac{1}{q}\left(\left(1+\frac{1}{q}\right)^2\ln 2+\frac{1}{q^2}[2e^q\mathcal{Ei}(-q)-e^{2q}\mathcal{Ei}(-2q)-\ln\gamma q]\right)\right\},$$

where[1]

$$q=\frac{ah}{k}, \quad \ln\gamma=0.5772, \quad -\mathcal{Ei}(-q)=\int_q^\infty \frac{e^{-t}}{t}dt.$$

It may be verified that the average temperatures of the preceding section are obtained by letting $q\to\infty$. Again, to see how sensitive $\bar T$ is to the precise distribution of the withdrawn heat $g(x)$ we compare $\bar T_C$ and $\bar T_L$ for the same total heat flow through the face $y=0$. The result is that $\bar T_L/\bar T_C\to 1$ as $q\to 0$, while from the preceding section we have this ratio $\to 1.15$ as $q\to\infty$. For q less than 1 the ratio does not deviate from 1 by more than 2 percent.

[1] For properties and tabulated values of the function $\mathcal{Ei}$ see Jahnke and Emde, *Tables of Functions* (Dover Publications, New York, 1943), p. 1.

We turn now to the flow $F(y)=-hT$ into the surface $x=0$. It appears that this function cannot be expressed in terms of known tabulated functions, but we can obtain some information about its maximum. For a general $g(x)$, we have from (11) and (13),

$$F(y)=\frac{2}{\pi}\int_0^\infty \frac{e^{-\alpha y}}{1+p^2\alpha^2}d\alpha\int_0^\infty g(t)(\sin\alpha t+\alpha p\cos\alpha t)dt.$$

From the known integrals

$$\frac{p^2\alpha}{1+\alpha^2p^2}=\int_0^\infty e^{-t/p}\sin\alpha t\,dt,\quad \frac{p}{1+\alpha^2p^2}=\int_0^\infty e^{-t/p}\cos\alpha t\,dt,$$

we obtain for $y>0$,

$$\begin{aligned}F(y)&=\frac{2}{\pi p}\int_0^\infty g(t)dt\int_0^\infty e^{-u/p}du\int_0^\infty \sin\alpha(t+u)e^{-\alpha y}d\alpha\\&=\frac{2}{\pi p}\int_0^\infty g(t)dt\int_0^\infty \frac{(u+t)e^{-u/p}du}{y^2+(u+t)^2}\\&=\frac{2}{\pi}\int_0^\infty g(t)dt\int_0^\infty \frac{(pu+t)e^{-u}du}{y^2+(pu+t)^2}.\end{aligned}$$

Hence,

$$\frac{dF}{dy}=-\frac{4y}{\pi}\int_0^\infty g(t)dt\int_0^\infty \frac{(pu+t)e^{-u}du}{[y^2+(pu+t)^2]^2}.$$

Thus, if g is non-negative, then dF/dy is negative for $y>0$ and F is decreasing. It follows that the maximum of F is given by

$$\max F=\lim_{y\to 0}F=\frac{2}{\pi}\int_0^\infty g(t)dt\int_0^\infty \frac{e^{-u}du}{pu+t}.$$

If we apply this result to the two cases (9) we obtain

$$\max F_C=\frac{2G}{\pi}[\ln\gamma q-e^q\mathcal{E}i(-q)],$$

$$\max F_L=\frac{2Aa}{\pi}\left[1-e^q\mathcal{E}i(-q)-\frac{1}{q}\{\ln\gamma q-e^q\mathcal{E}i(-q)\}\right].$$

A comparison of these two maxima for the same total flow shows $\max F_C/\max F_L=1.35,\ 1.32,\ 1.26$ for $q=1,$ 0.6, 0.3. The ratio tends to 1 as q tends to 0.

Inversion Formula

From Eq. (12) we are led to the integral equation

$$\int_0^\infty S(\alpha)(a\sin\alpha x+b\alpha\cos\alpha x)d\alpha=R(x),$$

where a, b are positive constants, $R(x)$ is a given function, and $S(\alpha)$ is to be solved for. Let

$$W(x)=\int_0^\infty S(\alpha)\sin\alpha x\,d\alpha.$$

The integral equation becomes

$$b\frac{dW}{dx}+aW=R(x),\quad W(0)=0,$$

whose solution is

$$W(x)=\frac{1}{b}e^{-ax/b}\int_0^x R(t)e^{at/b}dt.$$

Using the Fourier integral theorem,

$$\begin{aligned}S(\alpha)&=\frac{2}{\pi}\int_0^\infty W(x)\sin\alpha x\,dx\\&=\frac{2}{\pi b}\int_0^\infty \sin\alpha x e^{-ax/b}dx\int_0^x R(t)e^{at/b}dt\\&=\frac{2}{\pi b}\int_0^\infty R(t)e^{at/b}dt\int_t^\infty \sin\alpha x e^{-ax/b}dx.\end{aligned}$$

Thus

$$S(\alpha)=\frac{2}{\pi}\int_0^\infty R(t)\frac{a\sin\alpha t+b\alpha\cos\alpha t}{a^2+b^2\alpha^2}dt. \qquad (14)$$

This is the desired solution.[2] For $a=1$, $b=0$ we obtain the usual sine transform; for $a=0$, $b=1$, the cosine transform.

[2] The referee was kind enough to point out that formula (14) appears on p. 210 of R. V. Churchill, *Modern Operational Mathematics in Engineering*, (McGraw-Hill Book Company, Inc., New York, 1944). The derivation given there is different from the one presented here and makes use of the Laplace transform.

Oxygen-Free Single Crystals of Lead Telluride, Selenide, and Sulfide

W. D. LAWSON
Telecommunications Research Establishment, Ministry of Supply, Malvern, England
(Received December 3, 1951)

Anomalies in the type of conductivity of artificially prepared crystals of lead selenide and lead telluride have been found to be the result of the presence of oxygen. A method has been developed for eliminating oxygen from the specimens by melting the elements into the crucible in an atmosphere of hydrogen. It has also been found possible to grow single crystals of lead sulfide when oxygen is eliminated from the specimen in this way.

INTRODUCTION

A RECENT paper[1] has described a method of growing single crystals of lead telluride and lead selenide. A number of crystals were grown in which the composition of the original mixture of the elements varied gradually from 2 percent excess lead to 2 percent excess tellurium or selenium. These were grown for the purpose of testing the effect of these stoichiometric variations on conductivity and Hall constant. It had been expected that crystals with excess Pb would be *N*-type, i.e., conduct electricity mainly by electrons, and those with a deficit of Pb would be *P*-type, i.e., positive hole conductors.

This was found to be so for lead selenide crystals but the lead telluride crystals, surprisingly, were found to be all *P*-type with nearly the same conductivity, and it did not seem possible to grow *N*-type crystals of PbTe by this method even when as much as 5 percent excess lead was introduced.

A review of the stages leading up to the production of the final crystal, i.e., purity of initial materials, preparation of specimens and crucibles, growing and cooling conditions, revealed no difference capable of accounting for the discrepancy between selenide and telluride or for the failure to grow *N*-type PbTe. It did reveal however that the main impurity in both types of crystal was likely to be oxygen from residual and adsorbed air left in the crucible, in spite of lengthy pumping out. Accordingly an apparatus was constructed which would ensure the elimination of oxygen from the crucible and contents, by outgassing the crucible and melting in the elements in an atmosphere of hydrogen. It is shown in Fig. 1. The arms of the *Y*-tube are of Pyrex glass and the silica crucible is sealed to the bottom via a graded seal.

PREPARATION OF SPECIMEN

The apparatus, with weighed quantities of lead and tellurium or selenium inserted and held in position in separate arms of the *Y* by dimples in the glass, is connected to the pumping system and to a supply of dry oxygen-free hydrogen. The air is pumped out of the *Y*-tube and crucible, with the hydrogen inlet tap shut, and then the glass work is flushed several times with hydrogen.

The hydrogen is then turned on and lit at the burner. While hydrogen is flowing through the apparatus the crucible is first heated to red heat to expel adsorbed oxygen by combination, and then the lead is melted into it. The lead is also heated to red heat to insure complete reduction of any oxide. At this stage the hydrogen stream is turned off and the apparatus connected to the pumping system while the lead is still molten. This removes dissolved hydrogen from the lead. The tellurium or selenium is then melted into the crucible on top of the lead with the hydrogen flowing again. A small quantity—about 0.25 percent—is lost by evaporation. The final stage consists of pumping out the hydrogen and sealing off the crucible. This should be done rapidly before there is appreciable diffusion of air from the pumping system back into the crucible.

It is now essential to mix the elements thoroughly before growing a crystal; otherwise the crystal will not be homogeneous. This is achieved by heating the lead-selenium mixture to a temperature of about 350–400°C

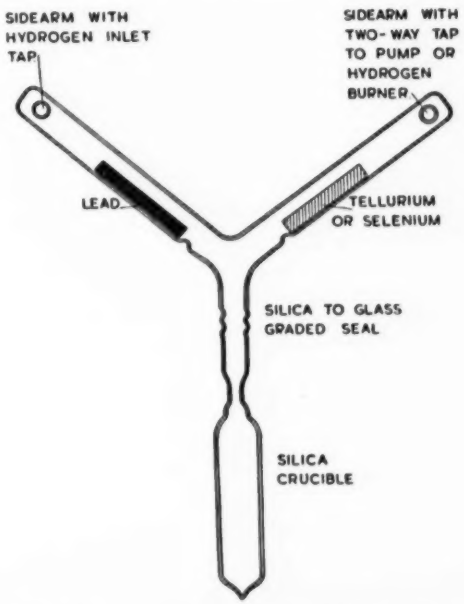

SIDEARM WITH HYDROGEN INLET TAP
SIDEARM WITH TWO-WAY TAP TO PUMP OR HYDROGEN BURNER
LEAD
TELLURIUM OR SELENIUM
SILICA TO GLASS GRADED SEAL
SILICA CRUCIBLE

FIG. 1. Apparatus for preparing specimens in an atmosphere of hydrogen.

[1] W. D. Lawson, J. Appl. Phys. **22**, 1444 (1951).

for about two days with the crucible horizontal. At this temperature lead selenide forms slowly and the molten elements mix with each other if the furnace is rocked. It is important to have complete combination into PbSe before further heating in the crystal growing furnace. The reason for this is that free selenium boils at 688°C. If an attempt is made to reach the melting point of lead selenide (1065°C) with any appreciable quantity of free selenium present the crucible will explode. This precaution is not necessary in the lead-tellurium mixture as the boiling point of free tellurium is 1390°C whereas PbTe melts at 904°C. The PbTe mixture is therefore heated at about 950°C for about two days to insure complete combination.

The specimens are now ready for crystal growing in the usual way.

It has been found possible to grow *N*-type crystals of lead telluride by preparing specimens with excess lead in this way. Specimens with excess tellurium come out *P*-type as expected. The lead selenide crystals prepared in this way also come out as expected.

It appears that the difference between the selenide and telluride lies in the fact that the telluride crystal absorbs oxygen as an impurity much more readily than the selenide crystal. Thus in the original experiments the presence of residual oxygen was making practically no difference to the selenide crystals, whereas it was being absorbed by the telluride lattice to an extent sufficient to mask the effect of excess lead. Some confirmation of this conclusion is available from other evidence. Experiments on the rectifying properties of these crystals have revealed that an oxide film forms easily on the telluride crystal, even at room temperature, whereas the selenide crystal has to be heated to about 600°C before an appreciable oxide layer forms.

LEAD SULFIDE CRYSTALS

Attempts to grow lead sulfide crystals from a melt in the manner described for lead telluride in reference 1 had been unsuccessful. When the sulfide was heated in a sealed crucible the crucible invariably exploded at or near the melting point, and when it was melted in an open crucible in an atmosphere of H_2S the resulting crystals were invariably full of blowholes. It was thought that these explosions and blowholes were caused by the release of free sulfur by dissociation at the melting point of PbS, 1114°C. At this temperature free sulfur must be a gas since its critical temperature is 1040°C.

An experiment in which the specimen was prepared by the hydrogen process just described revealed that it is possible to grow single crystals of lead sulfide provided the mixture is free from oxygen. A number of crystals have been grown in this way and here again excess lead gives *N*-type and excess sulfur *P*-type crystals. It is necessary here, as with selenide, to insure that all the sulfur is combined before heating above the boiling point of sulfur.

It is likely that the formation of SO_2 was causing explosions and blowholes in previous experiments.

ACKNOWLEDGMENTS

This paper is published by permission of the Chief Scientist, British Ministry of Supply. British Crown Copyright reserved. Reproduced by permission of the Controller of H. M. Stationery Office, England.

Letters to the Editor

The Separation of Cold-Work Distortion and Particle Size Broadening in X-Ray Patterns*

B. E. WARREN AND B. L. AVERBACH
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received January 28, 1952)

FOR a cubic crystal it is always possible to choose orthogonal axes so as to make any reflection have the indices (00l). It has been shown,[1] that for either distortion broadening or particle size broadening, the shape of the diffraction peak can be represented by a Fourier series.

$$P_{2\theta}=K\sum_{-\infty}^{+\infty} A_n(l)\cos 2\pi n h_3, \tag{1}$$

where $h_3=2a_3\sin\theta/\lambda$

for distortion $A_n{}^D(l)=\langle\cos 2\pi l Z_n\rangle_{Av}$,

for particle size $A_n{}^P=\frac{1}{N}\sum_{i=|n|+1}^{\infty}(i-|n|)n_i$.

If both types of broadening are present, the measured coefficient is the product of the coefficients for each effect.[2]

$$A_n(l)=A_n{}^P A_n{}^D(l). \tag{2}$$

If measurements are made for several orders of (00l), the two effects can be separated, since the distortion coefficient $A_n{}^D(l)$ depends upon the order, while the particle size coefficient $A_n{}^P$ is independent of l. For small l, and small values of n where Z_n is also small, $\langle\cos 2\pi l Z_n\rangle$ can be approximated by an exponential.

$$A_n{}^D(l)=\langle\cos 2\pi l Z_n\rangle_{Av}\rightarrow\exp[-2\pi^2 l^2\langle Z_n{}^2\rangle_{Av}].$$

Writing (2) in logarithmic form

$$\ln A_n(l)=\ln A_n{}^P-2\pi^2 l^2\langle Z_n{}^2\rangle_{Av}. \tag{3}$$

For several orders of a particular set of planes, a plot of $\ln A_n(l)$ *vs* l^2 gives a curve which is a straight line for the small values of l. The intercept on the axis of ordinates gives $\ln A_n{}^P$, and the slope gives $-2\pi^2\langle Z_n{}^2\rangle_{Av}$ from which the mean square strain can be computed. This method of separating the two effects is completely general in that it does not involve any assumption about the shape of the peaks. For orders of the general plane (hkl) of a cubic crystal, it is convenient to express $A_n{}^D(l)$ as $\exp[-2\pi^2 l_0{}^2\langle\Delta L^2\rangle/a^2]$, where $l_0{}^2=h^2+k^2+l^2$, $L=na_3$ and $\Delta L=a_3 Z_n$.

Figure 1 is a plot of $\ln A_n(l_0)$ *vs* $l_0{}^2$ for the first five orders of (111), measured on a single crystal of Cu−2 percent Si which has been rolled on the 110 plane in the $\bar{1}12$ direction to 50 percent reduction. The fact that the curves do not extrapolate to $\ln A_n(l_0)=0.0$, indicates that a small part of the broadening is due to particle size. The values of $A_n{}^P$ obtained from the intercepts are plotted against $L=na_3$ in Fig. 2. The negative slope of this line is the reciprocal of the average particle dimension perpendicular to the reflecting planes. The value obtained, $\bar{L}=1000$A, indicates that in the deforming of the crystal by rolling, the major part of the broadening is due to distortion, and the broadening by particle size is small but not small enough to neglect.

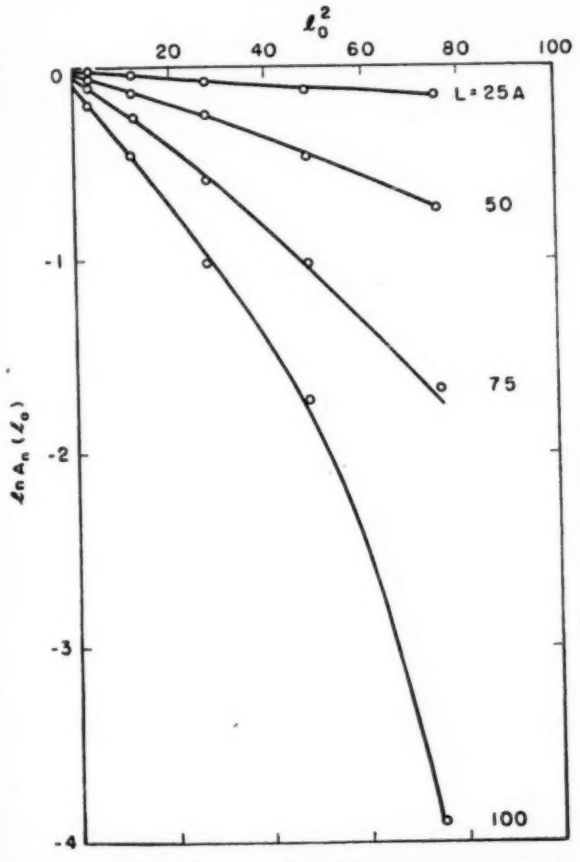


FIG. 1. Plot of $\ln A_n(l_0)$ *vs* $l_0{}^2$ for first five orders of (111), measured on a single crystal of Cu−2 percent Si which has been rolled to 50 percent reduction. $l_0{}^2=h^2+k^2+l^2$ and n is expressed in terms of $L=na_3$.

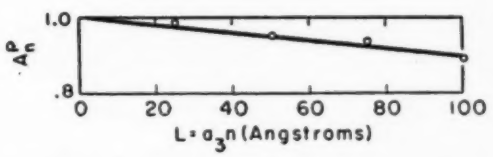


FIG. 2. Plot of $A_n{}^P$ *vs* L for single crystal of Cu−2 percent Si rolled to 50 percent reduction. The slope gives an average particle dimension $\bar{L}=1000$A.

* Research sponsored by the AEC under Contracts AT(30-1)-858 and AT(30-1)1002.
[1] B. E. Warren and B. L. Averbach, J. Appl. Phys. **21**, 595 (1950).
[2] This fact has been pointed out independently by A. J. C. Wilson and E. W. Hart.

Current-Carrying Wire Loops in a Simple Inhomogeneous Region

JAMES R. WAIT*
Newmont Exploration Limited, Jerome, Arizona
(Received January 17, 1952)

THE electromagnetic fields of a small current-carrying insulated wire loop situated in the plane of separation between two dissipative media will be calculated. The loop will be situated at the origin of cylindrical coordinates (ρ, φ, z) with the axis of the loop oriented in the z direction. The plane of separation is at $z=0$.

The formal solution for this problem has been given by Sommerfeld[1] for the case of a harmonically time-varying current. His result can be conveniently written in terms of electric vector potentials F_1 and F_2, as defined by Schelkunoff,[2] with a z component only, for the upper and lower regions, respectively. The upper and lower media are characterized by electrical constants σ_1, ϵ_1, μ_0 and σ_2, ϵ_2, μ_0, respectively. The time factor $\exp(i\omega t)$ is to be employed throughout. The mks system of units is used. The potentials are

$$F_1=\frac{I\delta A i\mu_0\omega}{2\pi}\int_0^{\infty}\frac{\lambda J_0(\lambda\rho)\exp(-u_2 z)}{u_1+u_2}d\lambda \tag{1}$$

and

$$F_2=\frac{I\delta A i\mu_0\omega}{2\pi}\int_0^{\infty}\frac{\lambda J_0(\lambda\rho)\exp(-u_1 z)}{u_1+u_2}d\lambda \tag{2}$$

where

$$u_1=(\lambda^2+\gamma_1^2)^{\frac{1}{2}},\quad u_2=(\lambda^2+\gamma_2^2)^{\frac{1}{2}},$$
$$\gamma_1^2=i\sigma_1\mu_0\omega-\epsilon_1\mu_0\omega^2$$

and

$$\gamma_2^2=i\sigma_2\mu_0\omega-\epsilon_2\mu_0\omega^2.$$

The current in the loop of area δA is given by I. The fields are then related to the potentials F_i by the following:

$$E_\varphi=\frac{\partial F_i}{\partial\rho},\quad i\mu_0\omega H_z=\left(-\gamma_i^2+\frac{\partial^2}{\partial z^2}\right)F_i,$$
$$i\mu_0\omega H_\rho=\frac{\partial^2 F_i}{\partial\rho\partial z},\quad\text{and}\quad H_\varphi=E_\rho=E_z=0. \tag{3}$$

The subscript i here takes the value 1 or 2 to pertain to the upper or lower region, respectively.

The integral involved in the expressions for F_1 and F_2 will be evaluated for the case when $z=0$. The common vector potential is now designated by F,

$$F=\frac{I\delta A i\mu_0\omega}{2\pi}\int_0^\infty\frac{\lambda J_0(\lambda\rho)}{u_1+u_2}d\lambda. \tag{4}$$

On multiplying the numerator and denominator by (u_1-u_2) the expression for F becomes

$$F=\frac{i\mu_0\omega I\delta A}{2\pi}\left[\int_0^\infty\frac{\lambda u_1 J_0(\lambda\rho)d\lambda}{(\gamma_1^2-\gamma_2^2)}-\int_0^\infty\frac{\lambda u_2 J_0(\lambda\rho)d\lambda}{(\gamma_1^2-\gamma_2^2)}\right]. \tag{5}$$

The numerator and denominator of the first expression is multiplied by $\lambda+u_1$ and in the second expression by $\lambda+u_2$. The result is

$$F=\frac{i\mu_0\omega I\delta A}{2\pi(\gamma_1^2-\gamma_2^2)}\left[\gamma_1^2\int_0^\infty\frac{J_0(\lambda\rho)\lambda d\lambda}{\lambda+u_1}-\gamma_2^2\int_0^\infty\frac{J_0(\lambda\rho)\lambda d\lambda}{\lambda+u_2}\right]. \tag{6}$$

These integrals have been evaluated by Foster[3] and are of the following type,

$$\int_0^\infty\frac{J_0(\lambda\rho)\lambda d\lambda}{\lambda+u_i}=\frac{1}{\gamma_i^2\rho^3}[1-(1+\gamma_i\rho)e^{-\gamma_i\rho}]. \tag{7}$$

Finally, the expression for F can be written,

$$F=\frac{i\mu_0\omega I\delta A}{2\pi(\gamma_1^2-\gamma_2^2)\rho^3}[(1+\gamma_1\rho)e^{-\gamma_1\rho}-(1+\gamma_2\rho)e^{-\gamma_2\rho}]. \tag{8}$$

The electric field component E_φ in the plane of separation (i.e., $z=0$) is then,

$$E_\varphi=\frac{\partial F}{\partial\rho}=\frac{i\mu_0\omega I\delta A}{2\pi(\gamma_1^2-\gamma_2^2)\rho^4}[(3+3\gamma_2\rho+\gamma_2^2\rho^2)e^{-\gamma_2\rho}$$
$$-(3+3\gamma_1\rho+\gamma_1^2\rho^2)e^{-\gamma_1\rho}]. \tag{9}$$

An expression for the vertical magnetic field component at the plane of separation can also be found. The field H_z is given everywhere by

$$i\mu_0\omega H_z=-\gamma_i^2F_i+\frac{\partial^2F_i}{\partial z^2}. \tag{10}$$

Now F_i satisfies the inhomogeneous wave equation in all the source free regions, that is

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial F_i}{\partial\rho}\right)-\gamma_i^2F_i+\frac{\partial^2F_i}{\partial z^2}=0 \tag{11}$$

for axial symmetry. The field H_z can then be written

$$i\mu_0\omega H_z=-\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial F_i}{\partial\rho}\right). \tag{12}$$

The differentiations of F can now be carried out to yield

$$H_z=\frac{I\delta A}{2\pi(\gamma_1^2-\gamma_2^2)\rho^5}[(9+9\gamma_1\rho+4\gamma_1^2\rho^2+\gamma_1^3\rho^3)e^{-\gamma_1\rho}$$
$$-(9+9\gamma_2\rho+4\gamma_2^2\rho^2+\gamma_2^3\rho^3)e^{-\gamma_2\rho}]. \tag{13}$$

The open-circuit voltage induced in a similar single-turn loop of area δB whose plane is also in the plane of separation is given by

$$e=-i\mu_0\omega\delta BH_z.$$

The mutual impedance between the loops is then given by

$$Z=e/I.$$

An expression is thus available which enables one to calculate the steady-state mutual impedance of two small wire loops contained in the plane of separation between two generally dissipative media. Such a physical situation could be represented by the earth's surface where the upper medium becomes nondissipative. Another situation might arise when the loops were lying on the floor of the ocean. Both media can then be highly dissipative.

The author appreciates the permission granted by Dr. A. A. Brant of Newmont Exploration Limited to publish this note.

* Research Engineer.

[1] A. Sommerfeld, Ann. Physik **81**, 1135 (1926).
[2] S. A. Schelkunoff, *Electromagnetic Waves* (D. Van Nostrand Company, Inc., New York, 1943), p. 128.
[3] R. M. Foster, Bell System Tech. J. **10**, No. 3, 408 (1931).

A Study of Oxide-Coated Cathode by X-Ray Diffraction Method

EISO YAMAKA
The Electrical Communication Laboratory, Tokyo, Japan
(Received October 29, 1951)

IN a preceding letter,[1] measurements of the crystal growth of (BaSr)O (Ba 50 percent, Sr 50 percent) through heat treatment by the x-ray diffraction method were reported.[2] Recently, by the same method, we have measured the crystal growth of BaO and (BaSrCa)O (Ba 47 percent, Sr 43 percent, Ca 10 percent). The results obtained are shown in Fig. 1. Each curve is a typical one of a few samples. In contrast with (BaSr)O, BaO grows quickly and above 1025°C, it reaches about one thousand angstroms. On the other hand, (BaSrCa)O grows slowly to about 1100°C; above 1100°C the line breadth of the diffracted x-ray beam becomes broader because of the preferential evaporation of BaO from the surface of the cathode and the decomposition of the Debye-Sherrer line of (BaSrCa)O into the lines of SrO and (BaSrCa)O.

The relation between the crystal size of BaO and its thermionic emission is shown in Fig. 2. To study the variation of emission due to the crystal growth only, pure electrolyt nickel sleeves were used. In this case heat activation will be dominant and impurity activation be small. Now, N_0, the number of activation centers, is given as follows in our case;

$$N_0(T)=C\exp(-E(L(T))/kT)=C\exp(-E(T)/kT),$$

where C is a constant and E is the activation energy of the impurity center. Kawamura *et al.*[3,4] concluded from their experiments that E is 4.4 ev for small crystals and 5.4 ev for large crystals. If

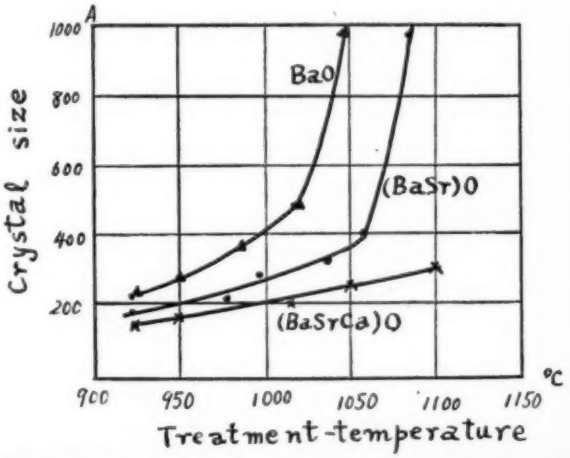


FIG. 1. Crystal size *vs* treatment-temperature in BaO, (BaSr)O(BaSrCa)O

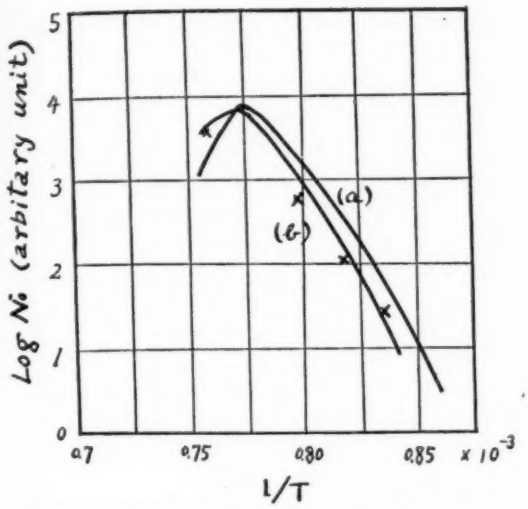


FIG. 2. Theoretical and experimental curves. (a) Theoretical curve. (b) Experimental curves.

we assume 4.4 ev for 200A and 5.4 ev for 1000A and moreover that E varies linearly with the crystal size, we obtain the calculated curve in Fig. 2. The experimental curve is derived from the Richardson curves taken in the temperature range where the work functions (this is not $E(T)$ above defined) are nearly the same, i.e., 1.9 ev. The coincidence of maximum points in Fig. 2 indicates a large value of E and shows the decrease of N_0 in the larger crystals. The relation between E and crystal size, however, cannot be deduced from the experimental curve, because the values of N_0 estimated from Richardson curves are not sufficiently accurate. We are now measuring N_0 directly using the method proposed by Kawamura *et al.*,[3] and the result will be reported later.

[1] E. Yamaka, J. Appl. Phys. **22**, 1087 (1951).
[2] E. Yamaka, J. Appl. Phys. **22**, 1037 (1951).
[3] H. Kawamura *et al.*, *Electron Emission and Semiconductor* (1947) (in Japanese), p. 245.
[4] H. Kawamura *et al.*, J. Phys. Soc. (Japan) **3**, 301 (1948).

Blue Glass Which Fades upon Nuclear Radiation

K. O. OTLEY AND W. A. WEYL
Department of Mineral Technology, The Pennsylvania State College, State College, Pennsylvania
(Received January 21, 1952)

DURING a systematic study of the effect of high energy radiation upon the light absorption of glasses[1] a type was found which is sensitive to x- and gamma-rays. This blue glass owes its color to the presence of elemental sulfur, probably in the form of S_2 molecules. Its absorption spectrum has a broad maximum between 5000 and 6000 a.u. Upon irradiation with x-rays the intensity of this absorption band decreases, a change which can be observed either with a visual or a photoelectric photometer.

Radiation sensitive glasses of this type can be obtained in various ways. It can be easily prepared on a borate basis by adding a few percent of elemental sulfur to fused sodium tetraborate. The resulting deep brown melt contains sodium polysulfides. The addition of increasing amounts of anhydrous boric oxide causes a change of its color from brown to gray, and finally, into a clear blue. The blue color center of this glass is assumed to be identical with that of the mineral lazurite or lapis lazuli or with that of artificial ultramarine. The color center probably[2] consists of S_2 molecules.

High energy radiation such as x-rays or the radiation of a cobalt source destroys the color center. Probably the trapping of one or more electrons by the S_2 molecule changes it into colorless sulfide ions.

[1] K. O. Otley, Glass Ind. **33**, No. 1, 24–27 (January, 1952).
[2] W. A. Weyl, *Coloured Glasses* (The Society of Glass Technology, Sheffield, England, 1951), p. 257.

The Effects of Lattice Imperfections on X-Ray Line Broadening

M. S. PATERSON
*The Institute for the Study of Metals, University of Chicago, Chicago, Illinois**
(Received December 31, 1951)

WARREN and Averbach[1] have shown that when x-ray powder pattern lines are broadened by lattice distortions or a small size effect in the crystals, the distribution of power in a line may be represented by

$$P'(2\theta)=KN\sum_{n=-\infty}^{\infty}A_n\cos 2\pi nh_3,$$

where K is a constant, N is the total number of unit cells in the sample, and h_3 is the reciprocal lattice coordinate normal to the reflecting planes, which is proportional to $\sin\theta/\lambda$. For pure distortion broadening, they show that $A_n=\langle\cos 2\pi lZ_n\rangle_{Av}$, where l is the value of h_3 at the intensity maximum and Z_n is the difference in displacements normal to the reflecting planes of two unit cells n spacings apart in this direction; for pure small-size broadening, A_n is the fraction of unit cells for which the position n spacings in the direction normal to the reflecting planes lies within the crystal. For distinction, we shall replace A_n by J_n in the former case and by V_n in the latter.

Now it is easily shown that when the two effects are present simultaneously,

$$P'(2\theta)=KN\,\Sigma_n\,V_nJ_n\cos 2\pi nh_3;$$

then the integral breadth of the line is

$$\beta=\frac{\lambda}{a_3\cos\theta}\cdot\frac{1}{\Sigma\,V_nJ_n},$$

or if the summation is approximated by an integral with continuous variable $t=na_3$,

$$\beta=\frac{\lambda}{\cos\theta}\cdot\frac{1}{\int V_tJ_t dt}.$$

This expression is identical with that derived by Stokes and Wilson[2,3] if V_0 and J_0 are taken as unity.

It will be seen more clearly that the quantity A_n for distortion broadening is equivalent to Stokes and Wilson's quantity J_t if in Warren and Averbach's calculation the effect of the displacement of a unit cell is expressed as a phase change in its F value, that is, we regard each unit cell as being situated at an undisturbed lattice site but having a phase factor in its F value which takes account of its actual displacement. Then, in terms of these F values,

$$A_n=\frac{1}{|F|^2}\cdot\overline{F_jF_{j+n}{}^*},$$

where F_j and F_{j+n} refer to pairs of unit cells n spacings apart in the direction normal to the reflecting planes, and the product $F_jF_{j+n}{}^*$ is averaged over all cells in the crystal.

The function J_t can be used to represent the line broadening effects of any type of imperfection in the lattice, including those involving more than one kind of atom (for example, imperfect ordering[2]). For example, in addition to the problem of internal strains, J_t has also been calculated for the case of stacking faults in hexagonal cobalt[4] and in deformed face-centered cubic crystals.[5] V_t always expresses the effect of the finite extent of the crystals.

If we write $\beta=\lambda/\epsilon\cos\theta$, in analogy with the small-size broadening from crystals of thickness ϵ, we have, in the absence of small-size effects, $\epsilon=\int J_t dt$, called by Wilson the "apparent particle size" (but likely to be dependent on θ and λ). J_t is a correlation coefficient for the scattering by unit cells distant t apart and ϵ is thus the integral breadth of the curve of this correlation function; that is, J_t is a measure of the degree of "coherence" or "correlation" in the lattice. In pure distortion broadening ϵ is the "average" distance in the lattice over which there is appreciable correlation of the phases of scattering, that is, coherence.

The use of quantities analogous to J_t is found in other branches of physics. In the study of turbulence,[6] for example, the correlation of the velocities at different points in the fluid is considered; in this case, the breadth of the correlation function is a measure of the average size of an eddy. More closely related to J_t is the Patterson function in x-ray crystallography;[7] this is a measure of the correlation of scattering at pairs of points a given distance apart within the individual unit cells, whereas J_t is concerned with the scattering from pairs of unit cells throughout the crystal.

* This work was done during leave of absence from Aeronautical Research Laboratories, Department of Supply, Melbourne, Australia.

[1] B. E. Warren and B. L. Averbach, J. Appl. Phys. **20**, 885 (1949); **21**, 595 (1950).
[2] A. J. C. Wilson, Proc. Roy. Soc. (London) **A181**, 360 (1943).
[3] A. R. Stokes and A. J. C. Wilson, Proc. Phys. Soc. (London) **56**, 174 (1944).
[4] A. J. C. Wilson, Proc. Roy. Soc. (London) **A180**, 277 (1942).
[5] M. S. Paterson (to be published).
[6] G. I. Taylor, Proc. Roy. Soc. (London) **A151**, 421 (1935).
[7] R. W. James, *The Optical Principles of the Diffraction of X-Rays* (London, 1948), p. 371.

The Transmission of X-Rays at Settings Near the Bragg Angle

G. N. RAMACHANDRAN
(*Department of Physics, Indian Institute of Science, Bangalore, India*)
(Received December 31, 1951)

IN a recent paper in this Journal, Campbell[1] has described some experiments illustrating how the intensity of the transmitted beam is increased at glancing angles of incidence (θ) close to the Bragg angle (θ_B). He concludes that the theory given by Laue[2] does not explain his results, "as his predicted intensity distributions are not the same as observed in this work." However, it must be noted that the curves shown in Laue's paper (which are probably what Campbell refers to as "predicted intensity distributions") relate to conditions which are very different from those involved in Campbell's experiments. The quantity $\exp(-\mu D/\gamma)$, where μ is the linear absorption coefficient, D the thickness, and γ the cosine of the angle made by the incident ray with the normal to the surface (which is the fraction of the incident intensity

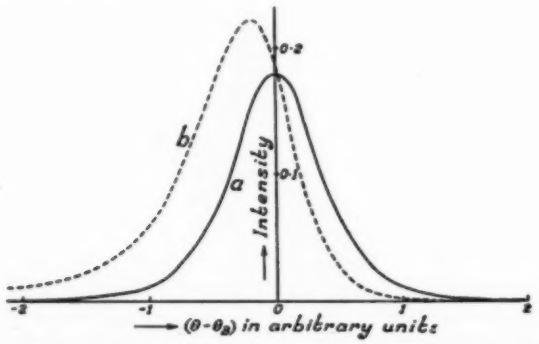


FIG. 1. Theoretical variation of intensity of (*a*) λ the reflected and (*b*) the transmitted beams for angles of incidence near the Bragg angle. Calcite $10\bar{1}1$ reflection with Cu K_α; thickness of crystal 0.4 mm. Note the relative displacement of the two maxima.

normally transmitted by the crystal at angles far away from θ_B), has a value of about 0.1 in Laue's calculations, while it is of the order of 10^{-3} to 10^{-6} in these experiments, so that the two are not strictly comparable. Laue's theory is, however, of more general validity and the present author has calculated, using it, the transmission and the reflection from crystals of calcite cut in the orientations employed by Campbell. The values of the real and imaginary parts of the $10\bar{1}1$ structure amplitude and other relevant are taken from the calculations of Zachariasen.[3] The curves have been calculated for two thicknesses, Fig. 1 for 0.40 mm and Fig. 2

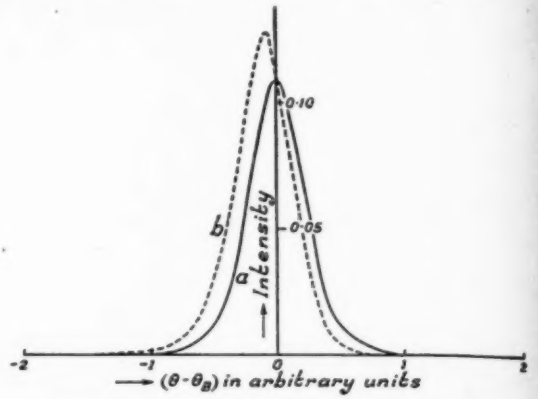


FIG. 2. Same as Fig. 1 for a crystal of thickness 1.0 mm.

for 1.00 mm which correspond respectively to Fig. 6 and Fig. 2 of Campbell's paper. The unit for the abscissa is the half-width of the reflection curve of a perfect nonabsorbing crystal, which is 7.2 seconds of arc. The ordinate is the intensity expressed as the fraction of the incident intensity. It would be noticed that the theory predicts sharp maxima both in the reflected and the transmitted beams near the Bragg angle as actually found by Campbell. However, the agreement between theory and experiment is not exact, because Figs. 1 and 2 only give the intensity distribution to be expected if the incident beam had an infinitely small divergence. In actual practice it would have a finite divergence (in the experiments mentioned above, about 15″) and the effect of this would be (*a*) to broaden the peak and (*b*) to depress the peak intensity. If this is taken into account, the agreement between theory and experiment must be considered good.

It is interesting to note that, according to theory, the peaks in the reflected and the transmitted beams are not coincident and that the separation between them should be less, the thicker the crystal, provided it is perfect throughout its thickness. A small relative displacement of the peaks is found in the curves plotted by Campbell, but he states that it is the result of a "slight misalignment of the crystal." It is not quite clear how such a misalignment would lead to a shift when two Geiger counters are used simultaneously, but it would be interesting to make a careful study in order to verify whether there is such a shift between the two maxima, as predicted by theory.

A more detailed discussion of the theory will be published elsewhere. It is found that the shape of the transmission curve depends on the thickness and absorption coefficient of the single crystal and that for a thin crystal the curve exhibits the "Helldunkelstruktur" observed by Borrmann,[4] while as the thickness increases, the bright peak becomes more and more prominent. For the crystals of quartz studied by Borrmann in his experiments, $\exp(-\mu D)$ was about 0.05, so that it is not surprising that Laue's curves fit his observations.

[1] H. N. Campbell, J. Appl. Phys. **22**, 1139 (1951).
[2] M. Von Laue, Acta Cryst. **2**, 106 (1949).
[3] W. H. Zachariasen, *The Theory of X-ray Diffraction in Crystals* (1945), p. 144.
[4] G. Borrmann, Physik. Z. **42**, 157 (1941).